
Cálculo III

Apuntes
Curso Código 525211

Dr. Raimund Bürger
Profesor Titular
Departamento de Ingeniería Matemática
& Centro de Investigación en Ingeniería Matemática (CI²MA)
Facultad de Ciencias Físicas y Matemáticas
Universidad de Concepción
Casilla 160-C
Concepción, Chile

15 de marzo de 2021

Índice general

Literatura	5
Capítulo 1. El espacio $\mathbb{R}^n$, espacios métricos y espacios vectoriales	7
1.1. $\mathbb{R}^n$ como espacio métrico	7
1.2. La convergencia en el espacio euclidiano $\mathbb{R}^n$	8
1.3. Los Teoremas de Bolzano-Weierstrass y de Heine-Borel para el espacio $\mathbb{R}^n$	10
1.4. Sucesiones de Cauchy, límites, y aplicaciones continuas en espacios métricos	13
1.5. Conjuntos conexos	15
1.6. Espacios vectoriales	18
1.7. Espacios vectoriales normados	20
1.8. La conexión entre el espacio $\mathbb{R}^n$ y el espacio vectorial $\mathbb{V}^n$	21
1.9. La equivalencia de normas en $\mathbb{V}^n$	22
1.10. Espacios de Hilbert	23
Capítulo 2. Funciones de $\mathbb{R}^n$ en $\mathbb{R}^m$ I: límites y continuidad, la diferencial, derivadas parciales y derivadas direccionales	27
2.1. Funciones reales de n variables	27
2.2. Límites direccionales y la continuidad direccional	27
2.3. Derivadas direccionales	29
2.4. Derivadas parciales	32
2.5. La diferenciabilidad en $\mathbb{R}^n$	34
2.6. Derivadas parciales de orden mayor	41
2.7. Aplicaciones de $\mathbb{R}^n$ en $\mathbb{R}^m$	49
2.8. La regla de la cadena	51
2.9. El Teorema del Valor Intermedio	55
Capítulo 3. Funciones de $\mathbb{R}^n$ en $\mathbb{R}^m$ II: aplicaciones	59
3.1. El Teorema de Taylor	59
3.2. Funciones implícitas	63
3.3. Funciones inversas	69
3.4. Extremos de funciones de varias variables	74
3.5. Extremos con restricciones y multiplicadores de Lagrange	81
Capítulo 4. Cálculo integral de funciones de varias variables I: teoría de la integración n -dimensional	85
4.1. Notación; sumas superiores e inferiores	85
4.2. Integrales de Riemann-Darboux inferiores y superiores	87

4.3.	La integral de Riemann para intervalos	88
4.4.	Sumas de Riemann	90
4.5.	Integrales iteradas	91
4.6.	Integrales de Riemann-Darboux inferiores y superiores sobre conjuntos acotados	96
4.7.	La integral de Riemann sobre conjuntos acotados	101
4.8.	La medida de conjuntos	104
4.9.	Conjuntos nulos especiales	110
4.10.	El principio de Cavalieri	111
4.11.	Los teoremas de valores intermedios del cálculo integral n -dimensional	113
4.12.	La integrabilidad sobre conjuntos generales	115
Capítulo 5.	La computación de integrales	117
5.1.	Dominios normales	117
5.2.	La regla de substitución para integrales n -dimensionales	123
5.3.	Centros de masa y momentos de inercia	132
Capítulo 6.	Análisis vectorial y teoremas integrales	135
6.1.	Curvas en $\mathbb{R}^n$ y el vector tangencial	135
6.2.	Funciones de variación acotada	139
6.3.	La longitud de una curva	144
6.4.	Campos vectoriales; la divergencia y el rotacional	149
6.5.	Integrales de línea	152
6.6.	Potenciales e independencia del camino de integrales de línea	154
6.7.	Superficies en $\mathbb{R}^3$	160
6.8.	Curvas en superficies, planos tangenciales y vectores normales	161
6.9.	El contenido de superficies e integrales de superficie	164
6.10.	El Teorema Integral de Green	169
6.11.	El Teorema Integral de Gauss	172
6.12.	El Teorema Integral de Stokes	175

Literatura

1. Marsden y Tromba, *Cálculo Vectorial*, 4a edición, Addison-Wesley/Iberoamericana, 1998.
2. Larson, Hostetler y Edwards, *Cálculo y Geometría Analítica*, Vol. 2. Mc Graw-Hill Interamericana, 1998.
3. Apostol, Calculus, Vol. I, II. Editorial Reverté, 1965.
4. Pita Ruiz. Cálculo Vectorial. Prentice-Hall Hispanoamericana, 1995.
5. Avello, Apuntes de Cálculo III, Universidad de Concepción, 1998.
6. Trench, Advanced Calculus. Harper/Row, 1978.
7. Anton, Cálculo y Geometría Analítica II. Limusa/Wesley, 1991.

Capítulo 1

El espacio $\mathbb{R}^n$, espacios métricos y espacios vectoriales

En este capítulo consideramos el espacio

$$\mathbb{R}^n = \{x \mid x = (x_1, \dots, x_n), \quad x_1, \dots, x_n \in \mathbb{R}\}, \quad n \in \mathbb{N},$$

cuyos elementos son *puntos* x que pueden ser descritos mediante un sistema de coordenadas.

1.1. $\mathbb{R}^n$ como espacio métrico

Primero definimos el concepto de la *distancia*, precisamente, la *distancia euclidiana* entre dos puntos de $\mathbb{R}^n$, para convertir $\mathbb{R}^n$ en un espacio métrico, donde recordamos la siguiente definición.

Definición 1.1. *Un conjunto $M \neq \emptyset$ se llama espacio métrico si para cada par de elementos $x, y \in M$ existe un número $d(x, y) \in \mathbb{R}$ tal que:*

- i) *Para todo $x, y \in M$, $d(x, y) \geq 0$, y $d(x, y) = 0$ si y sólo si $x = y$ (no-negatividad).*
- ii) *Para todo $x, y \in M$, $d(x, y) = d(y, x)$ (simetría).*
- iii) *Para todo $x, y, z \in M$, $d(x, y) \leq d(x, z) + d(z, y)$ (desigualdad triangular).*

Consideremos primero el caso $n = 2$ del plano con el sistema de coordenadas cartesiano y dos puntos $x = (x_1, x_2)$ e $y = (y_1, y_2)$. Según el Teorema de Pitágoras de la geometría euclidiana, la distancia entre x e y es dada por

$$d(x, y) = \sqrt{(y_1 - x_1)^2 + (y_2 - x_2)^2}.$$

Análogamente, para dos puntos $x = (x_1, \dots, x_n), y = (y_1, \dots, y_n) \in \mathbb{R}^n$ proponemos la siguiente función de distancia:

$$d(x, y) = \sqrt{\sum_{i=1}^n (y_i - x_i)^2}. \quad (1.1)$$

Teorema 1.1. *El espacio $\mathbb{R}^n$, equipado por la función $d : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ definida por (1.1), es un espacio métrico.*

Demostración. Basta verificar que la función d satisface los axiomas (i)–(iii) de la Definición 1.1. Es trivial verificar (i) y (ii); para verificar (iii), recordamos que

$$\sqrt{\sum_{i=1}^n (y_i - x_i)^2} \leq \sqrt{\sum_{i=1}^n (y_i - z_i)^2} + \sqrt{\sum_{i=1}^n (z_i - x_i)^2}$$

es el enunciado de la desigualdad de Minkowski. ■

Definición 1.2. La métrica d definida por (1.1) sobre $\mathbb{R}^n$ se llama métrica euclidiana sobre $\mathbb{R}^n$. En espacio $\mathbb{R}^n = (\mathbb{R}^n, d)$ se llama espacio euclidiano n -dimensional.

Queremos considerar las ε -vecindades

$$U_\varepsilon(x^0) := \{x \mid x \in \mathbb{R}^n, d(x, x^0) < \varepsilon\}.$$

de un punto $x^0 = (x_1^0, \dots, x_n^0)$ para $n = 1, 2, 3$. En el caso $n = 1$, $U_\varepsilon(x^0)$ es el intervalo abierto $(x^0 - \varepsilon, x^0 + \varepsilon)$. Para $n = 2$, $U_\varepsilon(x^0)$ es el interior del disco con centro x^0 y radio ε . Para $n = 3$, $U_\varepsilon(x^0)$ es el interior de la bola con centro x^0 y radio ε .

Definición 1.3.

1. Sea $-\infty < a_i < b_i < \infty$ para $i = 1, \dots, n$. El conjunto

$$[a, b] := \{x \mid x = (x_1, \dots, x_n), a_i \leq x_i \leq b_i, i = 1, \dots, n\}$$

se llama un intervalo cerrado de $\mathbb{R}^n$.

2. Sea $-\infty \leq a_i < b_i \leq \infty$ para $i = 1, \dots, n$. El conjunto

$$(a, b) := \{x \mid x = (x_1, \dots, x_n), a_i < x_i < b_i, i = 1, \dots, n\}$$

se llama un intervalo abierto de $\mathbb{R}^n$.

Teorema 1.2. Con respecto a la métrica euclidiana, cada intervalo (a, b) es abierto y cada intervalo $[a, b]$ es cerrado.

Demostración. Tarea. ■

Definición 1.4. Sean $-\infty < a_i < b_i < \infty$ para $i = 1, \dots, n$ e $I = [a, b]$ o $I = (a, b)$ el intervalo correspondiente cerrado o abierto. En este caso

$$\delta(I) = \sqrt{\sum_{i=1}^n (b_i - a_i)^2}$$

se llama el diámetro de I .

Definición 1.5. Un subconjunto $X \subset \mathbb{R}^n$ se llama acotado si existen un número $R > 0$ y un punto $x^0 \in \mathbb{R}^n$ tales que $X \subset U_R(x^0)$.

Teorema 1.3. Sea $X \subset \mathbb{R}^n$. Entonces los siguientes enunciados son equivalentes.

1. El conjunto X es acotado.
2. Existe un intervalo cerrado I tal que $X \subset I$.
3. Existe un $k \in \mathbb{N}$ tal que $X \subset U_k((0, \dots, 0))$.

Demostración. Tarea. ■

1.2. La convergencia en el espacio euclidiano $\mathbb{R}^n$

Recordemos primero el concepto de convergencia en un espacio métrico.

Definición 1.6. Sea (M, d) un espacio métrico. Entonces se dice que una sucesión $\{x_n\}_{n \in \mathbb{N}}$ de elementos de M converge al límite $x \in M$ si para cada $\varepsilon > 0$ existe un número N_ε tal que $d(x_n, x) < \varepsilon$ para todo $n > N_\varepsilon$.

Mediante el siguiente teorema podemos reducir la convergencia de sucesiones en $\mathbb{R}^n$ a la convergencia de las n sucesiones dadas por cada una de sus coordenadas.

Teorema 1.4. *Sea una sucesión en $\mathbb{R}^n$ dada por*

$$x^k = (x_1^k, \dots, x_n^k), \quad k \in \mathbb{N};$$

además sea $x = (x_1, \dots, x_n)$. Entonces los dos siguientes enunciados son equivalentes:

1. $\lim_{k \rightarrow \infty} x^k = x$,
2. $\lim_{k \rightarrow \infty} x_i^k = x_i$ para cada $i = 1, \dots, n$.

Demostración.

1. Supongamos primero que $x^k \rightarrow x$, es decir,

$$d(x^k, x) = \sqrt{\sum_{j=1}^n (x_j^k - x_j)^2} \rightarrow 0$$

para $k \rightarrow \infty$. Entonces, para cada $\varepsilon > 0$ existe un K_ε tal que

$$\forall k > K_\varepsilon : \quad \sum_{j=1}^n (x_j^k - x_j)^2 < \varepsilon^2.$$

Esto significa que para cada $i = 1, \dots, n$,

$$\forall k > K_\varepsilon : \quad |x_i^k - x_i| \leq \sqrt{\sum_{j=1}^n (x_j^k - x_j)^2} < \varepsilon.$$

Luego para cada $i = 1, \dots, n$ tenemos que $x_i^k \rightarrow x_i$ para $k \rightarrow \infty$.

2. Ahora supongamos que $x_i^k \rightarrow x_i$ para $k \rightarrow \infty$ y cada $i = 1, \dots, n$. Entonces para $\varepsilon > 0$ dado, existe un número $K_\varepsilon(i)$ para cada $i = 1, \dots, n$ tal que

$$\forall k > K_\varepsilon(i) : \quad |x_i^k - x_i| < \frac{\varepsilon}{\sqrt{n}},$$

de donde concluimos que

$$\begin{aligned} \forall k > \max\{K_\varepsilon(1), \dots, K_\varepsilon(n)\} : \quad d(x^k, x) &= \sqrt{\sum_{j=1}^n (x_j^k - x_j)^2} < \sqrt{\sum_{i=1}^n \left(\frac{\varepsilon}{\sqrt{n}}\right)^2} \\ &= \varepsilon, \end{aligned}$$

lo que significa que $x^k \rightarrow x$ cuando $k \rightarrow \infty$.

■

1.3. Los Teoremas de Bolzano-Weierstrass y de Heine-Borel para el espacio $\mathbb{R}^n$

Definición 1.7. Sea (M, d) un espacio métrico y $X \subset M$.

1. Un punto $h \in M$ se llama punto de acumulación de X si en cada vecindad de h existe un número infinito de elementos de X .
2. Definimos

$$H(X) := \{h \mid h \text{ es punto de acumulación de } X\}.$$

Comentamos que un punto de acumulación de un conjunto no necesariamente debe ser elemento del conjunto.

Ejemplo 1.1.

1. Consideremos para $M = \mathbb{R}$ el subconjunto $X \subset M$ dado por

$$X = \{x \mid x \in \mathbb{Q}, 0 < x < 1\} \cup \{2\}.$$

Entonces $H(X) = [0, 1]$. El conjunto de los puntos de acumulación no es contable.

2. Para $M = \mathbb{R}$, sea $X = \{1, \frac{1}{2}, \frac{1}{3}, \dots\}$. Aquí $H(X) = \{0\}$.

Teorema 1.5. Sea $X \subset M$. Entonces un punto $h \in M$ es un punto de acumulación de X si en cada vecindad de h existe un elemento $x \in X$ tal que $x \neq h$.

Demostración. Tarea. ■

Frecuentemente se utiliza el enunciado del siguiente teorema como definición de un punto de acumulación.

Teorema 1.6. Sea $X \subset M$. Un punto $h \in M$ es un punto de acumulación de X si y sólo si existe una sucesión $\{x_n\}_{n \in \mathbb{N}}$ tal que $x_n \in X$ y $x_n \neq h$ para $n \in \mathbb{N}$, y $x_n \rightarrow h$.

Demostración. Tarea. ■

Teorema 1.7. Un conjunto $X \subset M$ es cerrado si y sólo si $H(X) \subset X$.

Demostración. Tarea. ■

Teorema 1.8. Sea $X \subset M$ y X compacto. Entonces cada conjunto infinito $X_0 \subset X$ posee por lo menos un punto de acumulación en X .

Demostración. Tarea. ■

Para la demostración del Teorema de Bolzano-Weierstrass en $\mathbb{R}^n$, tenemos que definir el punto de acumulación de una sucesión.

Definición 1.8. Sea (M, d) un espacio métrico y $\{x_n\} = \{x_n\}_{n \in \mathbb{N}}$ una sucesión con $x_n \in M$ para $n \in \mathbb{N}$.

1. Un punto $v \in M$ se llama punto de acumulación de la sucesión $\{x_n\}$ si en cada vecindad de v se encuentra un número infinito de elementos de la sucesión.
2. Definimos

$$V(\{x_n\}) := \{v \mid v \text{ es punto de acumulación de } \{x_n\}\}.$$

Hay que enfatizar que el punto de acumulación de una sucesión $\{x_n\}$, según la Definición 1.8, es diferente del punto de acumulación del conjunto

$$X(\{x_n\}) := \{x \mid x = x_n \text{ para algún } n \in \mathbb{N}\}.$$

La razón es que un punto de acumulación de una sucesión puede ser generado por la repetición infinita de un número, mientras que esto está excluido para el punto de acumulación de un conjunto.

Ejemplo 1.2. Sea $M = \mathbb{R}$, y consideremos la sucesión $\{x_n\}$ definida por

$$x_n = (-1)^n \quad \text{para } n \in \mathbb{N}.$$

Aquí $X(\{x_n\}) = \{-1, 1\}$ y $V(\{x_n\}) = \{-1, 1\}$, pero ningún de los puntos de acumulación de la sucesión $\{x_n\}$ es un punto de acumulación del conjunto $X(\{x_n\})$, dado que este conjunto consiste en solamente dos elementos.

Teorema 1.9. Una sucesión acotada $\{x_n\}$, $x_n \in \mathbb{R}$ para todo $n \in \mathbb{N}$ posee por lo menos un punto de acumulación.

Demostración. Tarea. ■

Teorema 1.10 (Teorema de Bolzano-Weierstrass para $\mathbb{R}^n$). Cada conjunto infinito y acotado $X \subset \mathbb{R}^n$ posee por lo menos un punto de acumulación.

Demostración. Según el Teorema 1.6, $h \in \mathbb{R}^n$ es un punto de acumulación de X si y sólo si existe una sucesión $\{\xi^k\} = \{\xi^k\}_{k \in \mathbb{N}}$ tal que $\xi^k \in X$ y $\xi^k \neq h$ para todo $k \in \mathbb{N}$ y $\xi^k \rightarrow h$ cuando $k \rightarrow \infty$. Construiremos ahora una tal sucesión y un tal elemento h .

Dado que X tiene un número infinito de elementos, existe una sucesión $\{x_k\}_{k \in \mathbb{N}}$ tal que $x^k = (x_1^k, \dots, x_n^k) \in X$, donde $x^k \neq x^j$ si $k \neq j$. Aquí para cada $i = 1, \dots, n$ la sucesión $\{x_i^k\}_{k \in \mathbb{N}}$ es acotada. Concluimos que en virtud del Teorema 1.9, existen una subsucesión

$$\{x_1^{k_l^{(1)}}\}_{l \in \mathbb{N}} \subset \{x_1^k\}_{k \in \mathbb{N}}$$

y un número $h_1 \in \mathbb{R}$ tal que

$$\lim_{l \rightarrow \infty} x_1^{k_l^{(1)}} = h_1.$$

Dado que la sucesión $\{x_2^{k_l^{(1)}}\}_{l \in \mathbb{N}}$ es acotada, existen una subsucesión $\{k_l^{(2)}\}_{l \in \mathbb{N}}$ de $\{k_l^{(1)}\}_{l \in \mathbb{N}}$ y un número $h_2 \in \mathbb{R}$ tales que

$$\lim_{l \rightarrow \infty} x_2^{k_l^{(2)}} = h_2.$$

Continuando así, obtenemos en el n -ésimo paso que existen una subsucesión $\{k_l^{(n)}\}_{l \in \mathbb{N}}$ de $\{k_l^{(n-1)}\}_{l \in \mathbb{N}}$ y un número $h_n \in \mathbb{R}$ tales que

$$\lim_{l \rightarrow \infty} x_n^{k_l^{(n)}} = h_n.$$

Ahora elegimos $h = (h_1, \dots, h_n)$. Consideremos la sucesión $\{\xi^l\}_{l \in \mathbb{N}}$ definida por

$$\xi^l = \left(x_1^{k_l^{(n)}}, x_2^{k_l^{(n)}}, \dots, x_n^{k_l^{(n)}} \right).$$

Note que los elementos $\xi^l \in X$ satisfacen $\xi^{l_1} \neq \xi^{l_2}$ si $l_1 \neq l_2$. Según el Teorema 1.4,

$$\xi^l \xrightarrow{l \rightarrow \infty} (h_1, h_2, \dots, h_n) = h.$$

Puesto que $\xi^{l_0} = h$ puede ser válido para a lo más un l_0 , sabemos que h es un punto de acumulación de X . Esto concluye la demostración del Teorema de Bolzano-Weierstraß para $\mathbb{R}^n$. ■

A continuación trataremos la caracterización de los subconjuntos compactos de $\mathbb{R}^n$ demostrando la versión multi-dimensional del Teorema de Heine-Borel. Primeramente necesitamos el siguiente teorema.

Teorema 1.11. *Una sucesión $\{I_k\}_{k \in \mathbb{N}}$ de sub-intervalos de $\mathbb{R}^n$ tales que $I_{k+1} \subset I_k$ para $k \in \mathbb{N}$ tiene una intersección no vacía, es decir existe un x^0 tal que $x^0 \in I_k$ para todo $k \in \mathbb{N}$.*

Demostración. Definimos

$$I_k := [a^k, b^k] = \{x \mid x = (x_1, \dots, x_n), a_i^k \leq x_i \leq b_i^k, \quad i = 1, \dots, n\}.$$

1. Primero fijamos i . Puesto que $I_k \supset I_{k+1}$, sabemos que

$$a_i^k \leq a_i^{k+1} \leq b_i^{k+1} \leq b_i^k, \quad k \in \mathbb{N}.$$

Entonces la sucesión $\{a_i^k\}_{k \in \mathbb{N}}$ es monótona, creciente y acotada. Si definimos

$$x_i^0 = \lim_{k \rightarrow \infty} a_i^k,$$

se tiene que $a_i^k \leq x_i^0 \leq b_i^k$ para cada $i = 1, \dots, n$ y $k \in \mathbb{N}$.

2. Determinamos de esta manera x_i^0 para cada $i = 1, \dots, n$. Entonces el punto $x^0 = (x_1^0, \dots, x_n^0)$ está contenido en cada uno de los intervalos I_k para $k \in \mathbb{N}$, luego

$$x^0 \in \bigcap_{k=1}^{\infty} I_k. \quad \blacksquare$$

Con la ayuda de este teorema podemos demostrar ahora el Teorema de Heine-Borel.

Teorema 1.12 (Teorema de Heine-Borel). *Un subconjunto $X \subset \mathbb{R}^n$ es compacto si y sólo si es acotado y cerrado.*

Demostración.

1. Sea X compacto. Entonces X es cerrado y acotado (resultado del Cálculo I).
2. Sea X acotado y cerrado. Supongamos que X no fuera compacto. Entonces existe un cubrimiento abierto F de X que *no* contiene un sub-cubrimiento finito de X .
 - a) Dado que X es acotado, existe un intervalo cerrado I_1 tal que $X \subset I_1$. Subdividimos I_1 en 2^n sub-intervalos cerrados dividiendo los n intervalos de coordenadas en dos mitades. Para por lo menos uno de estos sub-intervalos, el cual llamaremos I_2 , el conjunto $X \cap I_2$ *no* puede ser cubierto por un sub-cubrimiento finito de F . Evidentemente, $\delta(I_2) = \delta(I_1)/2$. Continuando la subdivisión obtenemos una sucesión de intervalos cerrados I_k tal que $I_k \supset I_{k+1}$; además, I_k tiene la propiedad de que $X \cap I_k$ *no* puede ser cubierto por un sub-cubrimiento finito de F , y

finalmente $\delta(I_k) = \delta(I_1)/2^{k-1}$. Según el Teorema 1.10 existe un punto x^0 tal que $x^0 \in I_k$ para $k \in \mathbb{N}$. Dado que cada I_k contiene un número infinito de elementos de X , x^0 también es un punto de acumulación de X , y dado que X es cerrado, $x^0 \in X$.

- b) El cubrimiento F contiene por lo menos un conjunto abierto S tal que $x^0 \in S$. Puesto que S es abierto, existe un $\varepsilon > 0$ con $U_\varepsilon(x^0) \subset S$. Sea k tan grande que $\delta(I_k) = \delta(I_1)/2^{k-1} < \varepsilon$; en este caso se tiene que $I_k \subset U_\varepsilon(x^0) \subset S$, dado que para todo $x \in I_k$, $d(x, x^0) \leq \delta(I_k) < \varepsilon$. Entonces I_k y por lo tanto $X \cap I_k$ puede ser cubierto por un conjunto abierto de F , en contradicción con la construcción. ■

Los subconjuntos compactos de $\mathbb{R}^n$ también pueden ser caracterizados de la siguiente manera.

Teorema 1.13. *Sea $X \subset \mathbb{R}^n$ un conjunto no vacío. Entonces los siguientes enunciados son equivalentes:*

1. X es compacto.
2. Cada sucesión $\{x^k\}_{k \in \mathbb{N}}$ con $x_k \in X$ posee un punto de acumulación en X .
3. Cada subconjunto infinito de X posee un punto de acumulación en X .

Demostración. Tarea. ■

1.4. Sucesiones de Cauchy, límites, y aplicaciones continuas en espacios métricos

Definición 1.9. *Una sucesión $\{x_n\}$ en un espacio métrico (M, d) se llama sucesión de Cauchy si para cada $\varepsilon > 0$ existe un N_ε tal que para todo $m, n > N_\varepsilon$, $d(x_m, x_n) < \varepsilon$.*

Definición 1.10. *Un espacio métrico M se llama completo si cada sucesión de Cauchy converge en M , es decir, posee un límite en M .*

Teorema 1.14. *El espacio euclidiano $\mathbb{R}^n$ es completo.*

Demostración.

1. Sea $\{x^k\}_{k \in \mathbb{N}}$ con $x^k = (x_1^k, \dots, x_n^k)$ una sucesión de Cauchy en $\mathbb{R}^n$. Entonces para cada $\varepsilon > 0$ existe un N_ε tal que

$$\forall k, l > N_\varepsilon : \quad d(x^k, x^l) = \sqrt{\sum_{j=1}^n (x_j^k - x_j^l)^2} < \varepsilon.$$

2. Sea i fijo ($i = 1, \dots, n$). Entonces

$$\forall k, l > N_\varepsilon : \quad |x_i^k - x_i^l| \leq \sqrt{\sum_{j=1}^n (x_j^k - x_j^l)^2} < \varepsilon,$$

es decir $\{x_i^k\}_{k \in \mathbb{N}}$ es una sucesión de Cauchy en $\mathbb{R}$. Puesto que $\mathbb{R}$ es completo, $\{x_i^k\}_{k \in \mathbb{N}}$ converge a un límite $x_i \in \mathbb{R}$.

3. Para $x = (x_1, \dots, x_n)$ obtenemos $x^k \rightarrow x$, según el Teorema 1.4.

■

Definición 1.11. Sean M_1 y M_2 espacios métricos, $f : M_1 \rightarrow M_2$. Sea x_0 un punto de acumulación del dominio $D(f)$ de f , y supongamos que existe un $g \in M_2$ tal que para cada sucesión $\{x_n\}_{n \in \mathbb{N}} \subset D(f)$ con $x_n \neq x_0$ y $x_n \rightarrow x_0$ se tiene que $\lim_{n \rightarrow \infty} f(x_n) = g$. En este caso, g se llama el límite de f en x_0 . Escribimos

$$g = \lim_{x \rightarrow x_0} f(x).$$

Notamos que la aplicación f no necesariamente debe ser definida en x_0 .

Teorema 1.15. Sean M_1 y M_2 espacios métricos y $f : M_1 \rightarrow M_2$. Entonces f es continua en un punto de acumulación x_0 de $D(f)$ si y sólo si

$$\lim_{x \rightarrow x_0} f(x) = f(x_0).$$

Demostración. Tarea. ■

En particular, este teorema es válido para cada punto interior x_0 de $D(f)$.

Si (M, d) es un espacio métrico y $M_0 \neq \emptyset$, $M_0 \subset M$, entonces (M_0, d) también es un espacio métrico. De vez en cuando es útil interpretar un subconjunto $M_0 \subset M$ como un propio espacio métrico.

Ahora sean (M_1, d_1) y (M_2, d_2) espacios métricos y $f : M_1 \rightarrow M_2$. Si $X_2 \subset M_2$, la *imagen recíproca* por f de X_2 es el conjunto

$$f^{-1}(X_2) := \{x_1 \in M_1 \mid f(x_1) \in X_2\}.$$

Tenemos la siguiente caracterización de la continuidad de f sobre M_1 .

Teorema 1.16. Sean (M_1, d_1) y (M_2, d_2) espacios métricos y f una aplicación de M_1 en M_2 . Entonces los tres siguientes enunciados son equivalentes.

1. La aplicación f es continua sobre M_1 .
2. Para cada conjunto abierto $X_2 \subset M_2$, el conjunto $f^{-1}(X_2)$ es abierto en M_1 .
3. Para cada conjunto cerrado $Y_2 \subset M_2$, el conjunto $f^{-1}(Y_2)$ es cerrado en M_1 .

Demostración.

1. Demostramos primero la implicación $(1) \Rightarrow (2)$. Sea X_2 un conjunto abierto en M_2 . Consideremos su imagen recíproca $f^{-1}(X_2) =: X_1 \subset M_1$. Sea x_0 un punto arbitrario en X_1 ; entonces $f(x_0) \in X_2$, y dado que X_2 es abierto, existe un $\varepsilon > 0$ tal que

$$U_\varepsilon(f(x_0)) = \{x_2 \in M_2 \mid d_2(f(x_0), x_2) < \varepsilon\} \subset X_2.$$

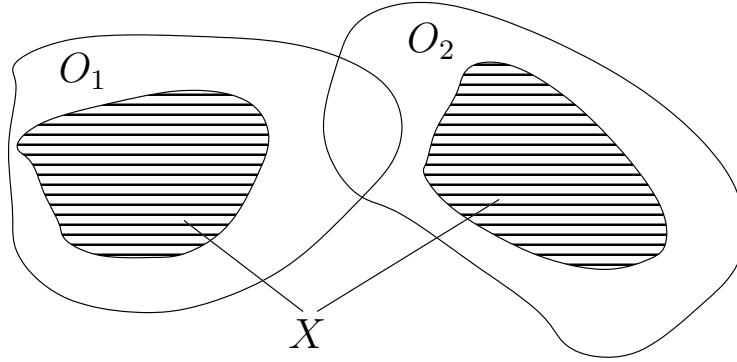
Dado que f es continua en x_0 , existe un $\delta_\varepsilon > 0$ tal que para todo $x \in M_1$ con $d_1(x_0, x) < \delta_\varepsilon$ se tiene que $d_2(f(x_0), f(x)) < \varepsilon$. Entonces se tiene que

$$\forall x \in U_{\delta_\varepsilon}(x_0) = \{x_1 \in M_1 \mid d_1(x_0, x_1) < \delta_\varepsilon\} : f(x) \in U_\varepsilon(f(x_0)) \subset X_2.$$

Pero esto significa que

$$U_{\delta_\varepsilon}(x_0) \subset f^{-1}(X_2),$$

entonces $f^{-1}(X_2)$ es abierto.

FIGURA 1.1. Un conjunto X disconexo

2. Demostramos ahora la implicación $(2) \Rightarrow (3)$. Sea Y_2 un conjunto cerrado en M_2 . Entonces su complemento $\mathbb{C}Y_2 = M_2 \setminus Y_2$ es abierto en M_2 . Según (2), $f^{-1}(\mathbb{C}Y_2)$ es abierto en M_1 . Pero

$$f^{-1}(\mathbb{C}Y_2) = M_1 \setminus f^{-1}(Y_2) = \mathbb{C}(f^{-1}(Y_2)).$$

Esto implica que $f^{-1}(Y_2)$ es cerrado en M_1 .

3. Demostramos finalmente la implicación $(3) \Rightarrow (1)$. Supongamos que existe un $\bar{x} \in M_1$ donde f no es continua. Entonces existen un $\bar{\varepsilon} > 0$ y una sucesión $\{x_n\} \in M_1$ y

$$\lim_{n \rightarrow \infty} x_n = \bar{x} \quad (1.2)$$

tal que $d_2(f(\bar{x}), f(x_n)) \geq \bar{\varepsilon}$ para todo $n \in \mathbb{N}$. Entonces se debe tener que

$$f(x_n) \in M_2 \setminus U_{\bar{\varepsilon}}(f(\bar{x})) = \mathbb{C}[U_{\bar{\varepsilon}}(f(\bar{x}))].$$

Dado que $U_{\bar{\varepsilon}}(f(\bar{x}))$ es abierto en M_2 , $\mathbb{C}[U_{\bar{\varepsilon}}(f(\bar{x}))]$ es cerrado en M_2 . Para todo $n \in \mathbb{N}$ ahora se tiene que

$$x_n \in f^{-1}(\mathbb{C}[U_{\bar{\varepsilon}}(f(\bar{x}))]),$$

y este conjunto es cerrado en M_1 . Pero en virtud de (1.2), también se debe tener que

$$\bar{x} \in f^{-1}(\mathbb{C}[U_{\bar{\varepsilon}}(f(\bar{x}))]),$$

lo que contradice $d(f(\bar{x}), f(\bar{x})) \geq \bar{\varepsilon} > 0$. ■

1.5. Conjuntos conexos

Definición 1.12. Sea (M, d) un espacio métrico y $X \neq \emptyset$ un subconjunto de M .

1. El conjunto X se llama disconexo si existen conjuntos abiertos O_1 y O_2 en M tales que

$$X \subset O_1 \cup O_2; \quad O_1 \cap X \neq \emptyset; \quad O_2 \cap X \neq \emptyset; \quad X \cap O_1 \cap O_2 = \emptyset. \quad (1.3)$$

2. El conjunto X se llama conexo si no es disconexo.

Evidentemente, un conjunto X es desconexo si puede ser descompuesto en dos componentes disjuntos a través de los conjuntos O_1 y O_2 .

Ejemplo 1.3. El conjunto $X = [0, 1) \cup (1, 2]$ es desconexo en $\mathbb{R}$. Consideremos los conjuntos abiertos $O_1 = (-1, 1)$ y $O_2 = (1, 3)$, entonces $X \subset O_1 \cup O_2$; $O_1 \cap X = [0, 1) \neq \emptyset$, $O_2 \cap X = (1, 2] \neq \emptyset$, pero $O_1 \cap O_2 = \emptyset$ y por lo tanto, también $X \cap O_1 \cap O_2 = \emptyset$.

Ejemplo 1.4. Sea X un subconjunto de un espacio métrico (M, d) que consiste en solamente un punto, por ejemplo $X = \{\xi\}$. En este caso, X es conexo. Para ver esto, supongamos que X fuera desconexo. En este caso, existen conjuntos abiertos O_1 y O_2 en M tales que $X \subset O_1 \cup O_2$ y $O_1 \cap X \neq \emptyset$, entonces $\xi \in O_1 \cap X$, y $O_2 \cap X \neq \emptyset$, entonces $\xi \in O_2 \cap X$. Por lo tanto, $\xi \in X \cap O_1 \cap O_2$, es decir, $X \cap O_1 \cap O_2 \neq \emptyset$, una contradicción.

Es fácil caracterizar los subconjuntos conexos de $\mathbb{R}$.

Teorema 1.17. Un conjunto $X \subset \mathbb{R}$ es conexo si y sólo si X contiene sólo un elemento o X es un intervalo.

Demostración. Si X contiene sólo un elemento, entonces X es trivialmente conexo según Ejemplo 1.4. Podemos suponer entonces que X contiene por lo menos dos puntos diferentes.

1. Sea X conexo. Supongamos que X no fuera un intervalo. Entonces existen puntos a , b y c tales que $a < b < c$ y $a \in X$, $c \in X$, pero $b \notin X$. Los conjuntos $O_1 = (-\infty, b)$ y $O_2 = (b, \infty)$ son conjuntos abiertos que satisfacen

$$X \subset O_1 \cup O_2; \quad O_1 \cap X \neq \emptyset; \quad O_2 \cap X \neq \emptyset; \quad X \cap O_1 \cap O_2 = \emptyset,$$

lo que implica que X es desconexo. ¡Contradicción!

2. Sea X un intervalo. Supongamos que X fuera desconexo. Entonces existen conjuntos abiertos O_1 y O_2 tales que

$$X \subset O_1 \cup O_2; \quad O_1 \cap X \neq \emptyset; \quad O_2 \cap X \neq \emptyset; \quad X \cap O_1 \cap O_2 = \emptyset.$$

Sea $a \in O_1 \cap X$, $b \in O_2 \cap X$ y sin restricción de generalidad $a < b$. Demostraremos que $(a, b) \not\subset X$ y por lo tanto X no puede ser un intervalo. Si tuviéramos que $(a, b) \subset X$, también $[a, b] \subset X$. Para los conjuntos $O_1 \cap [a, b]$ y $O_2 \cap [a, b]$ sabemos que

$$O_1 \cap [a, b] = (\mathbb{C}O_2) \cap [a, b], \quad O_2 \cap [a, b] = (\mathbb{C}O_1) \cap [a, b],$$

así que ambos conjuntos son cerrados. Podemos entonces concluir que

$$\xi = \sup(O_1 \cap [a, b]) \in O_1 \cap [a, b].$$

Aquí se debe cumplir $\xi \neq b$ (sino tendríamos que $b \in X \cap O_1 \cap O_2$). Según la definición de ξ , ninguna vecindad de ξ puede enteramente pertenecer a $O_1 \cap [a, b]$, así que ξ debe ser un punto de acumulación de $O_2 \cap [a, b]$. Puesto que $O_2 \cap [a, b]$ es cerrado, $\xi \in O_2 \cap [a, b]$, y por lo tanto

$$\xi \in O_1 \cap O_2 \cap [a, b] \subset O_1 \cap O_2 \cap X,$$

que contradice a $O_1 \cap O_2 \cap X = \emptyset$.

■

Teorema 1.18. Sean (M_1, d_1) y (M_2, d_2) espacios métricos. Consideremos una aplicación $f : M_1 \rightarrow M_2$ continua sobre el conjunto conexo $X \subset D(f) \subset M_1$. Entonces la imagen de X por f ,

$$f(X) = \{y \mid y = f(x), x \in X\}$$

es un conjunto conexo en M_2 .

Demostración. Supongamos que $f(X)$ fuera disconexo. Entonces existen conjuntos abiertos $O_1, O_2 \subset M$ tales que

$$f(X) \subset O_1 \cup O_2; \quad f(X) \cap O_1 \neq \emptyset; \quad f(X) \cap O_2 \neq \emptyset; \quad f(X) \cap O_1 \cap O_2 = \emptyset.$$

1. Consideremos los siguientes conjuntos:

$$\tilde{O}_i := \{x \in X \mid f(x) \in O_i\} \subset M_1, \quad i = 1, 2.$$

Ambos conjuntos son no vacíos. Si $x \in \tilde{O}_i \cap X$, entonces $y = f(x) \in O_i \cap f(X)$, y dado que O_i es abierto, existe un $\varepsilon(x) > 0$ tal que

$$U_{\varepsilon(x)}(f(x)) \subset O_i.$$

Como f es continua en x , existe un $\delta(x) > 0$ tal que

$$d_2(f(x), f(\xi)) < \varepsilon(x) \quad \text{para todo } \xi \in X \text{ con } d_1(x, \xi) < \delta(x),$$

lo que significa que

$$f(U_{\delta(x)} \cap X) \subset U_{\varepsilon(x)}(f(x)) \cap f(X).$$

Ahora definimos

$$O_i^* := \bigcup_{x \in \tilde{O}_i} U_{\delta(x)}(x).$$

Ambos conjuntos son abiertos, y

$$f(O_i^* \cap X) \subset O_i \cap f(X).$$

2. Analizaremos ahora las propiedades de los conjuntos abiertos O_1^* y O_2^* .

a) Para ver que

$$X \subset O_1^* \cup O_2^*,$$

consideremos $x \in X$; evidentemente, o $f(x) \in O_1$ o $f(x) \in O_2$ y por lo tanto o $x \in \tilde{O}_1$ o $x \in \tilde{O}_2$. El enunciado sigue de $\tilde{O}_i \subset O_i^*$.

b) Para ver que

$$X \cap O_i^* \neq \emptyset \quad \text{para } i = 1, 2,$$

nos acordamos de que debido a $f(X) \cap O_i \neq \emptyset$ existe un $\tilde{y}_i \in f(X) \cap O_i$. Por lo tanto, existe un $\tilde{x}_i \in \tilde{O}_i$ con $f(\tilde{x}_i) = \tilde{y}_i$. Para este $\tilde{x}_i$ se tiene que $\tilde{x}_i \in X \cap O_i^*$.

c) Se tiene que

$$X \cap O_1^* \cap O_2^* = \emptyset,$$

porque si existiera $x \in X$ tal que $x \in O_1^*$ y $x \in O_2^*$, tendríamos $f(x) \in O_1$ y $f(x) \in O_2$, en contradicción con $f(X) \cap O_1 \cap O_2 = \emptyset$.

De (a)–(c) se tiene que el conjunto X es desconexo, una contradicción. ■

Teorema 1.19 (Teorema del valor intermedio). *Sea (M, d) un espacio métrico y sea la función $f : M \rightarrow \mathbb{R}$ continua sobre el conjunto conexo $X \subset D(f) \subset M$. Sean $a, b \in X$ tales que $f(a) \leq f(b)$; además, sea $\eta \in \mathbb{R}$ tal que*

$$f(a) \leq \eta \leq f(b).$$

Entonces existe al menos un $\xi \in X$ con $f(\xi) = \eta$.

Demostración.

1. Si $f(a) = f(b)$, basta tomar $\xi = a$.
2. Sea $f(a) < f(b)$. Según el Teorema 1.18, $f(X)$ es conexo y debido a $f(X) \subset \mathbb{R}$, en virtud del Teorema 1.17, $f(X)$ es un intervalo tal que $[f(a), f(b)] \subset f(X)$. Entonces para cada $\eta \in [f(a), f(b)]$ existe a lo menos un $\xi \in X$ con $f(\xi) = \eta$. ■

1.6. Espacios vectoriales

Definición 1.13. *Sea K un cuerpo arbitrario. Entonces un conjunto $V \neq \emptyset$ se llama espacio vectorial sobre K si para cada $v, w \in V$, la suma $v + w$ y para cada $v \in V$ y $\lambda \in K$, el múltiple λv son definidos y nuevamente son elementos de V , tal que para vectores $u, v, w \in V$ arbitrarios y $\lambda, \mu \in K$ se cumplen las siguientes reglas:*

$$v + w = w + v, \tag{V1}$$

$$u + (v + w) = (u + v) + w, \tag{V2}$$

$$\exists 0 \in V : \forall v \in V : v + 0 = v, \tag{V3}$$

$$\forall v \in V : \exists -v \in V : v + (-v) = 0, \tag{V4}$$

$$\lambda(\mu v) = (\lambda\mu)v, \tag{V5}$$

$$\exists e \in K : ev = v, \tag{V6}$$

$$\lambda(v + w) = \lambda v + \lambda w, \tag{V7}$$

$$(\lambda + \mu)v = \lambda v + \mu v. \tag{V8}$$

Definición 1.14. *Sea V un espacio vectorial arbitrario sobre un cuerpo K , y sean $v_1, \dots, v_n \in V$ y $\lambda_1, \dots, \lambda_n \in K$.*

1. *El vector*

$$v = \sum_{i=1}^n \lambda_i v_i$$

se llama combinación lineal de los vectores $v_1, \dots, v_n$. Esta combinación se llama trivial si $\lambda_1 = \dots = \lambda_n = 0$; sino se llama no trivial.

2. Los vectores $v_1, \dots, v_n$ se llaman linealmente dependientes si existe una combinación lineal no trivial tal que

$$\sum_{i=1}^n \lambda_i v_i = 0.$$

En todos los demás casos se llaman linealmente independientes.

3. El sub-espacio lineal

$$L(v_1, \dots, v_n) := \left\{ v \mid v = \sum_{i=1}^n \lambda_i v_i, \lambda_1, \dots, \lambda_n \in K \right\}$$

es la envoltura lineal de los vectores $v_1, \dots, v_n$.

4. Si los vectores $v_1, \dots, v_n$ son linealmente independientes y $L(v_1, \dots, v_n) = V$, entonces $v_1, \dots, v_n$ es una base de V y n es la dimensión de V .

Para nuestro análisis el espacio vectorial más importante es el siguiente.

Definición 1.15. El conjunto

$$\mathbb{V}^n = \{ \vec{v} = \{v_1, \dots, v_n\} \mid v_1, \dots, v_n \in \mathbb{R} \}$$

con las operaciones

$$\vec{v} + \vec{w} = \{v_1 + w_1, \dots, v_n + w_n\}; \quad \lambda \vec{v} = \{\lambda v_1, \dots, \lambda v_n\} \quad (\lambda \in \mathbb{R})$$

se llama espacio vectorial n -dimensional sobre $\mathbb{R}$.

Se verifica fácilmente que efectivamente $\mathbb{V}^n$ es un espacio vectorial si definimos $\vec{0} = \{0, \dots, 0\}$.

Considerando los vectores linealmente independientes

$$\vec{e}_1 = \{1, 0, 0, \dots, 0\}, \quad \vec{e}_2 = \{0, 1, 0, \dots, 0\}, \quad \vec{e}_n = \{0, 0, \dots, 0, 1\},$$

podemos representar cada vector $\vec{v} = (v_1, \dots, v_n) \in \mathbb{V}^n$ en la forma

$$\vec{v} = \sum_{i=1}^n v_i \vec{e}_i.$$

Definición 1.16. Los vectores $\vec{e}_1, \dots, \vec{e}_n$ se llaman base natural de $\mathbb{V}^n$.

El cálculo hace frecuentemente uso de espacios vectoriales de funciones.

Definición 1.17. Sea X un conjunto arbitrario no vacío. El conjunto de todas las funciones

$$F(X) := \{ f \mid f : X \rightarrow \mathbb{R}, D(f) = X \}$$

con las operaciones

$$\begin{aligned} f + g : & \quad (f + g)(x) = f(x) + g(x), \quad x \in X, \\ \lambda f : & \quad (\lambda f)(x) = \lambda f(x), \quad \lambda \in \mathbb{R}, x \in X \end{aligned}$$

se llama espacio vectorial de las aplicaciones de X en $\mathbb{R}$.

Se verifican fácilmente los axiomas del espacio vectorial, considerando como elemento nulo la función que desaparece idénticamente.

Definición 1.18. Sea X un espacio métrico. El subconjunto de $F(X)$ de las funciones continuas,

$$C(X) := \{f \mid f : X \rightarrow \mathbb{R}, f \text{ es continua sobre } D(f) = X\}$$

se llama espacio vectorial de las funciones continuas de X en $\mathbb{R}$.

Aquí igualmente se verifica fácilmente la satisfacción de los axiomas de un espacio vectorial.

1.7. Espacios vectoriales normados

Entre los espacios vectoriales son importantes aquellos espacios que permiten la definición de la longitud de un vector.

Definición 1.19. Un espacio vectorial V sobre $\mathbb{R}$ se llama espacio vectorial normado si para cada $v \in V$ podemos definir un número $\|v\| \geq 0$ tal que para todo $v, w \in V$ y $\lambda \in \mathbb{R}$ se satisfacen los siguientes axiomas:

$$\|v\| = 0 \text{ si y sólo si } v = 0, \quad (\text{N1})$$

$$\|\lambda v\| = |\lambda| \|v\|, \quad (\text{N2})$$

$$\|v + w\| \leq \|v\| + \|w\|. \quad (\text{N3})$$

La cantidad $\|v\|$ se llama norma o longitud de v .

Ejemplo 1.5. Consideremos el espacio $\mathbb{V}^n$ y definamos para un vector $\vec{v} = \{v_1, \dots, v_n\}$ arbitrario

$$\|\vec{v}\| := \sqrt{\sum_{i=1}^n v_i^2}.$$

Evidentemente, $\|\vec{v}\| \geq 0$; también es fácil verificar (N1) y (N2). El axioma (N3) asume aquí la forma

$$\sqrt{\sum_{i=1}^n (v_i + w_i)^2} \leq \sqrt{\sum_{i=1}^n v_i^2} + \sqrt{\sum_{i=1}^n w_i^2},$$

pero esto es justamente la desigualdad de Minkowski.

Definición 1.20. La norma definida sobre el espacio vectorial n -dimensional $\mathbb{V}^n$ por

$$\|\vec{v}\| = \sqrt{\sum_{i=1}^n v_i^2}$$

se llama norma euclidiana, y el espacio $\mathbb{V}^n = (\mathbb{V}^n, \|\cdot\|)$ espacio vectorial euclidiano n -dimensional.

Ejemplo 1.6. Sea $[a, b]$ un intervalo cerrado y $C[a, b]$ el espacio de las funciones continuas sobre $[a, b]$. En este caso, podemos definir la norma del máximo

$$\|f\| := \max_{[a, b]} |f(x)|. \quad (1.4)$$

Teorema 1.20. *Sea V un espacio vectorial normado. Entonces*

$$d(v, w) = \|v - w\|$$

define una métrica sobre V .

Demostración. Tarea. ■

Esta definición de una métrica nos permite aplicar en un espacio vectorial todos los conceptos de un espacio métrico, es decir, la convergencia, vecindades, conjuntos abiertos, cerrados y compactos, puntos de acumulación, etc. Demostraremos esto con dos ejemplos.

Teorema 1.21. *Sea una sucesión en $\mathbb{V}^n$ dada por*

$$\vec{v}^k = \{v_1^k, \dots, v_n^k\}, \quad k \in \mathbb{N};$$

además, sea $\vec{v} = \{v_1, \dots, v_n\}$. Entonces los siguientes enunciados son equivalentes:

1. $\lim_{k \rightarrow \infty} \vec{v}^k = \vec{v}$.
2. $\lim_{k \rightarrow \infty} v_i^k = v_i$ para cada $i = 1, \dots, n$.

Demostración. La demostración es completamente análoga a la demostración del Teorema 1.4. ■

Teorema 1.22. *Sea $\{f_n\}_{n \in \mathbb{N}}$ una sucesión en $C[a, b]$, y sea $f \in C[a, b]$. Entonces los siguientes enunciados son equivalentes.*

1. $f_n \rightarrow f$ en $C[a, b]$.
2. $f_n(x) \rightarrow f(x)$ para cada $x \in [a, b]$.

Demostración. La equivalencia de los enunciados es evidente si tomamos en cuenta que el primer enunciado significa que

$$d(f_n, f) = \max_{x \in [a, b]} |f_n(x) - f(x)| \xrightarrow{n \rightarrow \infty} 0.$$
■

Interpretando una función como un “punto” en un espacio métrico podemos reducir el concepto complicado de la convergencia uniforme al concepto más simple de la convergencia de una sucesión de puntos en un espacio métrico.

1.8. La conexión entre el espacio $\mathbb{R}^n$ y el espacio vectorial $\mathbb{V}^n$

Los elementos de ambos espacios son definidos por n componentes reales, y las fórmulas de la distancia entre dos puntos de $\mathbb{R}^n$ y entre dos vectores de $\mathbb{V}^n$ son idénticas. En la literatura es muy común identificar ambos espacios. Pero es importante distinguir entre $\mathbb{R}^n$ como un espacio de *puntos* y $\mathbb{V}^n$ como un espacio de *vectores*.

Sin embargo, los vectores de $\mathbb{V}^n$ pueden ser utilizados para la descripción del espacio $\mathbb{R}^n$.

Definición 1.21. *Un vector $\vec{v} = \{v_1, \dots, v_n\} \in \mathbb{V}^n$ define una aplicación*

$$\vec{v} : \mathbb{R}^n \rightarrow \mathbb{R}^n, \quad D(\vec{v}) = \mathbb{R}^n,$$

$$x \mapsto \vec{v}(x) = \vec{v}((x_1, \dots, x_n)) = (x_1 + v_1, \dots, x_n + v_n).$$

Esta aplicación se llama traslación.

Ya observamos que existe una analogía fuerte entre las definiciones de los espacios $\mathbb{R}^n$ y $\mathbb{V}^n$.

1.9. La equivalencia de normas en $\mathbb{V}^n$

Definimos en la norma euclidiana para $\vec{v} = \{v_1, \dots, v_n\} \in \mathbb{V}^n$ por

$$\|\vec{v}\| = \|\vec{v}\|_2 = \sqrt{\sum_{i=1}^n v_i^2}.$$

Se verifica fácilmente que las expresiones

$$\|\vec{v}\|_1 := \sum_{i=1}^n |v_i| \quad (1.5)$$

y

$$\|\vec{v}\|_\infty := \max_{1 \leq i \leq n} |v_i| \quad (1.6)$$

igualmente definen normas en $\mathbb{V}^n$. Demostraremos aquí que es suficiente considerar la norma euclidiana en $\mathbb{V}^n$, puesto que todas las normas sobre $\mathbb{V}^n$ son equivalentes (en un sentido a precisar).

Teorema 1.23. *Consideramos el espacio métrico $\mathbb{V}^n$ con la norma euclidiana. Sea $\vec{v}_0 \in \mathbb{V}^n$ fijo. Entonces para cada $R > 0$, el conjunto*

$$S_R(\vec{v}_0) = \{\vec{v} \in \mathbb{V}^n \mid \|\vec{v} - \vec{v}_0\| = R\}$$

es compacto.

Demostración. Tarea. ■

Teorema 1.24. *Todas las normas en $\mathbb{V}^n$ son equivalentes en el siguiente sentido. Si $\|\cdot\|^*$ y $\|\cdot\|^{**}$ son normas en $\mathbb{V}^n$, entonces existen constantes positivas κ y K tales que*

$$\forall \vec{v} \in \mathbb{V}^n : \kappa \|\vec{v}\|^* \leq \|\vec{v}\|^{**} \leq K \|\vec{v}\|^*.$$

Demostración. Sea $\|\cdot\|$ una norma sobre $\mathbb{V}^n$ y $\|\cdot\|_2$ la norma euclidiana. Para demostrar el teorema es suficiente demostrar que existen constantes $\tilde{\kappa} > 0$ y $\tilde{K} > 0$ tales que

$$\forall \vec{v} \in \mathbb{V}^n : \tilde{\kappa} \|\vec{v}\|_2 \leq \|\vec{v}\| \leq \tilde{K} \|\vec{v}\|_2. \quad (1.7)$$

1. Sean $\vec{e}_1, \dots, \vec{e}_n$ los vectores de la base natural de $\mathbb{V}^n$. Entonces

$$\|\vec{v}\| = \left\| \sum_{i=1}^n v_i \vec{e}_i \right\| \leq \sum_{i=1}^n \|v_i \vec{e}_i\| = \sum_{i=1}^n |v_i| \|\vec{e}_i\| \leq \|\vec{v}\|_2 \sum_{i=1}^n \|\vec{e}_i\| = \tilde{K} \|\vec{v}\|_2,$$

donde definimos

$$\tilde{K} = \sum_{i=1}^n \|\vec{e}_i\|.$$

Esto concluye la demostración de la parte derecha de la desigualdad (1.7).

2. Consideremos ahora en el espacio métrico $\mathbb{V}^n$ con la norma euclidiana la aplicación

$$\mathbb{V}^n = D(f) \ni \vec{v} \mapsto f(\vec{v}) = \|\vec{v}\| \in \mathbb{R}.$$

Utilizando la desigualdad triangular y la primera parte de esta demostración, obtenemos para todo $\vec{v}, \vec{w} \in \mathbb{V}^n$

$$|f(\vec{v}) - f(\vec{w})| = |\|\vec{v}\| - \|\vec{w}\|| \leq \|\vec{v} - \vec{w}\| \leq \tilde{K}\|\vec{v} - \vec{w}\|_2 = \tilde{K}d(\vec{v}, \vec{w}),$$

lo que implica que f es continua sobre $\mathbb{V}^n$. Por otro lado, según el Teorema 1.23, el conjunto

$$S = \{\vec{v} \in \mathbb{V}^n \mid \|\vec{v}\|_2 = 1\}$$

es compacto. Dado que f es continua, f posee un mínimo absoluto sobre S , es decir, existe un $\vec{v}_0 \in S$ tal que

$$\forall \vec{v} \in S : \quad \tilde{\kappa} := f(\vec{v}_0) \leq f(\vec{v}),$$

es decir,

$$\forall \vec{v} \in S : \quad \tilde{\kappa} = \|\vec{v}_0\| \leq \|\vec{v}\|.$$

Dado que $\vec{v}_0 \in S$, o sea, $\|\vec{v}_0\|_2 = 1$, $\vec{v}_0 \neq \vec{0}$ y por lo tanto $\tilde{\kappa} > 0$.

Sea ahora $\vec{v} \in \mathbb{V}^n$ dado. Si $\vec{v} = \vec{0}$, trivialmente $\tilde{\kappa}\|\vec{v}\|_2 \leq \|\vec{v}\|$. Si $\vec{v} \neq \vec{0}$,

$$\frac{1}{\|\vec{v}\|_2} \vec{v} \in S$$

y por lo tanto

$$\tilde{m} \leq \left\| \frac{1}{\|\vec{v}\|_2} \vec{v} \right\| = \frac{\|\vec{v}\|}{\|\vec{v}\|_2},$$

es decir $\tilde{\kappa}\|\vec{v}\|_2 \leq \|\vec{v}\|$, lo que demuestra la parte izquierda de la desigualdad (1.7).

Esto concluye la demostración del teorema. ■

Comentamos que un resultado análogo al Teorema 1.24 no es válido en espacio normados generales. Por ejemplo el espacio $C[a, b]$ admite normas que poseen estructuras tan diferentes que no existe una equivalencia tal como en el Teorema 1.24.

1.10. Espacios de Hilbert

Definición 1.22. *Un espacio vectorial V sobre $\mathbb{R}$ se llama espacio de Hilbert o espacio unitario si para cada $v, w \in V$ está definido un número $(v, w) \in \mathbb{R}$ (el producto escalar o producto interior de los vectores v y w) tal que para todo $u, v, w \in V$ y $\lambda \in \mathbb{R}$ se cumplen las siguientes axiomas:*

$$(v + w, u) = (v, u) + (w, u), \tag{H1}$$

$$(\lambda v, w) = \lambda(v, w), \tag{H2}$$

$$(v, w) = (w, v), \tag{H3}$$

$$(v, v) \geq 0; \quad (v, v) = 0 \text{ si y sólo si } v = 0. \tag{H4}$$

Comentamos que en la literatura normalmente se dice “espacio de Hilbert” para un espacio vectorial *completo* con un producto escalar que satisface (H1)–(H4). Los espacios considerados aquí se llaman entonces “espacios pre-Hilbert”.

Ejemplo 1.7.

1. Para el espacio $\mathbb{V}^n$, $\vec{v} = (v_1, \dots, v_n)$ y $\vec{w} = (w_1, \dots, w_n)$ definimos

$$(\vec{v}, \vec{w}) = \sum_{i=1}^n v_i w_i.$$

Es fácil verificar que $(\cdot, \cdot)$ satisface (H1)–(H4) y entonces es un producto escalar. También escribimos para este producto escalar $(\vec{v}, \vec{w}) = \vec{v} \cdot \vec{w}$.

2. Para un intervalo cerrado $[a, b]$ consideramos el espacio $C[a, b]$ y definimos para $f, g \in C[a, b]$

$$(f, g) = \int_a^b f(x)g(x) dx.$$

Los axiomas (H1)–(H4) pueden ser verificados fácilmente.

3. Definimos el espacio de sucesiones

$$l^2 = \left\{ v = \{v_i\}_{i \in \mathbb{N}} \left| \sum_{i=1}^{\infty} v_i^2 < \infty \right. \right\}.$$

Para elementos arbitrarios $v = \{v_i\}_{i \in \mathbb{N}} \in l^2$, $w = \{w_i\}_{i \in \mathbb{N}} \in l^2$ definimos $v + w := \{v_i + w_i\}_{i \in \mathbb{N}}$ y $\lambda v = \{\lambda v_i\}_{i \in \mathbb{N}}$. En este caso, $v + w, \lambda v \in l^2$; además se verifica fácilmente que para $v, w \in l^2$ la series

$$\sum_{i=1}^{\infty} v_i w_i$$

convergen absolutamente, y que

$$(v, w) := \sum_{i=1}^{\infty} v_i w_i$$

define un producto escalar que convierte l^2 en un espacio de Hilbert.

Teorema 1.25 (Desigualdad de Schwarz). Si V es un espacio de Hilbert, entonces

$$\forall v, w \in V : \quad |(v, w)| \leq \sqrt{(v, v)} \cdot \sqrt{(w, w)}.$$

Demostración.

1. Si $v = 0$ o $w = 0$, la desigualdad es trivialmente + correcta.
2. Sea entonces $v \neq 0$ y $w \neq 0$. Para $\lambda \in \mathbb{R}$ consideremos la desigualdad

$$0 \leq (\lambda v + w, \lambda v + w) = \lambda^2(v, v) + 2\lambda(v, w) + (w, w).$$

Para

$$\lambda = -\frac{(v, w)}{(v, v)}$$

obtenemos

$$0 \leq \frac{(v, w)^2}{(v, v)} - 2 \frac{(v, w)^2}{(v, v)} + (w, w)$$

o equivalentemente,

$$(v, w)^2 \leq (v, v)(w, w),$$

lo que concluye la demostración. ■

Teorema 1.26. *Cada espacio de Hilbert es un espacio normado si definimos*

$$\|v\| = \sqrt{(v, v)},$$

Demostración. Tarea. ■

Capítulo 2

Funciones de $\mathbb{R}^n$ en $\mathbb{R}^m$ I: límites y continuidad, la diferencial, derivadas parciales y derivadas direccionales

2.1. Funciones reales de n variables

Empezaremos ahora el análisis de funciones $f : \mathbb{R}^n \rightarrow \mathbb{R}$. Dado que ambos $\mathbb{R}^n$ y $\mathbb{R}$ son espacios métricos, todas las definiciones y todos los teoremas del capítulo anterior que se refieren a espacios métricos quedan válidos.

Para el caso particular de una función

$$f : \mathbb{R}^2 \ni (x, y) \mapsto f(x, y) \in \mathbb{R}$$

todavía podemos visualizar la función considerando el conjunto de puntos

$$S_f := \{(x, y, z) \mid (x, y) \in D(f), z = f(x, y)\}.$$

Bajo hipótesis apropiadas, S_f es una *superficie* en $\mathbb{R}^3$.

Ejemplo 2.1.

1. Consideremos la función $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ con

$$D(f) = \{(x, y) \mid x^2 + y^2 < 1\}, \quad f(x, y) = \sqrt{1 - x^2 - y^2}.$$

Evidentemente, S_f es el hemisfério superior del radio 1.

2. Para $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ con

$$D(f) = \mathbb{R}^2, \quad f(x, y) = xy$$

(ver Figura 2.1), S_f es un paraboloide hiperbólico.

2.2. Límites direccionales y la continuidad direccional

Para el estudio de las funciones $f : \mathbb{R}^n \rightarrow \mathbb{R}$ necesitamos refinar los conceptos de la convergencia y de la continuidad.

Definición 2.1. Una dirección en $\mathbb{R}^n$ es dada por un vector $\vec{a} \in \mathbb{V}^n$ con $\|\vec{a}\| = 1$.

Queremos estudiar funciones en una dirección $\vec{a}$ dada. Entonces consideramos f en una vecindad de un punto interior x^0 de $D(f)$ solamente sobre la recta

$$G_{\vec{a}}(x^0) := \{x \mid x = x^0 + h\vec{a}, h \in \mathbb{R}\},$$

es decir estudiamos la función

$$\varphi(h) = f(x^0 + h\vec{a}) \quad (x^0 + h\vec{a} \in D(f))$$

del parámetro real de la recta h .

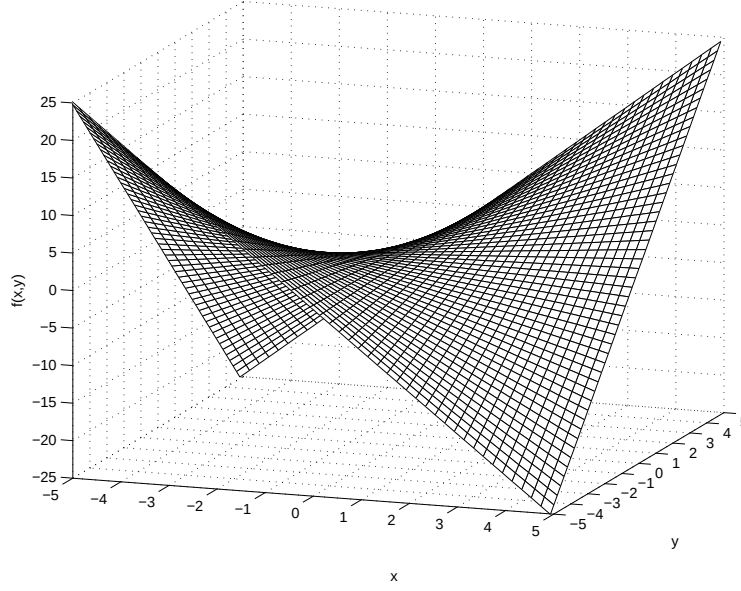


FIGURA 2.1. La función $f(x, y) = xy$ definida en el Ejemplo 2.1.

Definición 2.2. Sea $f : \mathbb{R}^n \rightarrow \mathbb{R}$. Se dice que en un punto interior $x^0 \in D(f)$, la función f posee el límite g en la dirección $\vec{a}$ si

$$\lim_{h \rightarrow 0} f(x^0 + h\vec{a}) = g.$$

Escribimos

$$\vec{a}\text{-}\lim_{x \rightarrow x^0} f(x) = g.$$

Evidentemente, si una función posee el límite g , entonces todos sus límites direccionales existen e igualan g . Se podría pensar que por otro lado, si todos estos límites direccionales existen e igualan g , entonces f posee el límite g en x^0 . Esto no es así, tal como demuestra el siguiente ejemplo.

Ejemplo 2.2. Sea la función $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ dada por

$$f(x, y) = \begin{cases} 0 & \text{si } (x, y) = (0, 0) \text{ o } y \neq x^2, \\ 1 & \text{si } y = x^2, x \neq 0. \end{cases}$$

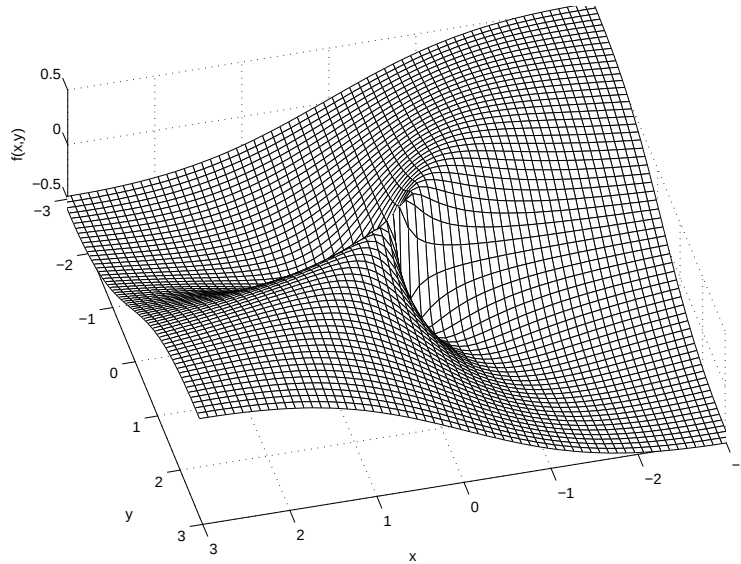
Evidentemente, $f(0, 0) = 0$ y para cada dirección $\vec{a}$,

$$\vec{a}\text{-}\lim_{(x,y) \rightarrow (0,0)} f(x, y) = 0.$$

Pero

$$\lim_{k \rightarrow \infty} f\left(\frac{1}{k}, \frac{1}{k^2}\right) = 1,$$

por lo tanto f no posee un límite en $(0, 0)$, y en particular es discontinua en $(0, 0)$.

FIGURA 2.2. La función $f(x, y)$ definida en el Ejemplo 2.3.

Definición 2.3. Una función $f : \mathbb{R}^n \rightarrow \mathbb{R}$ se llama continua en un punto interior x^0 de $D(f)$ en la dirección $\vec{a}$ si

$$\vec{a}\text{-}\lim_{x \rightarrow x^0} f(x) = f(x^0).$$

El Ejemplo 2.2 presenta una función que es discontinua en $(0, 0)$, pero que es continua en $(0, 0)$ en cada dirección, por lo tanto una función que es continua en un punto en cada dirección no necesariamente es continua en este punto.

Ejemplo 2.3. Se considera la función

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad f(x, y) = \begin{cases} \frac{xy}{x^2 + y^2} & \text{si } (x, y) \neq (0, 0), \\ 0 & \text{si } (x, y) = (0, 0). \end{cases}$$

Esta función no es continua en el punto $(0, 0)$ dado que

$$\lim_{k \rightarrow \infty} f\left(\frac{1}{k}, \frac{1}{k}\right) = \lim_{k \rightarrow \infty} \frac{1}{2} = \frac{1}{2} \neq f(0, 0).$$

Pero f es continua en las direcciones $\vec{e}_1$ y $\vec{e}_2$, dado que

$$\lim_{h \rightarrow 0} f(0 + h, 0) = 0; \quad \lim_{h \rightarrow 0} f(0, 0 + h) = 0.$$

2.3. Derivadas direccionales

Sea $f : \mathbb{R}^n \rightarrow \mathbb{R}$. Analizaremos primero el comportamiento de f en la vecindad de un punto interior x^0 de $D(f)$ en una dirección. Dado así f es una función de una variable,

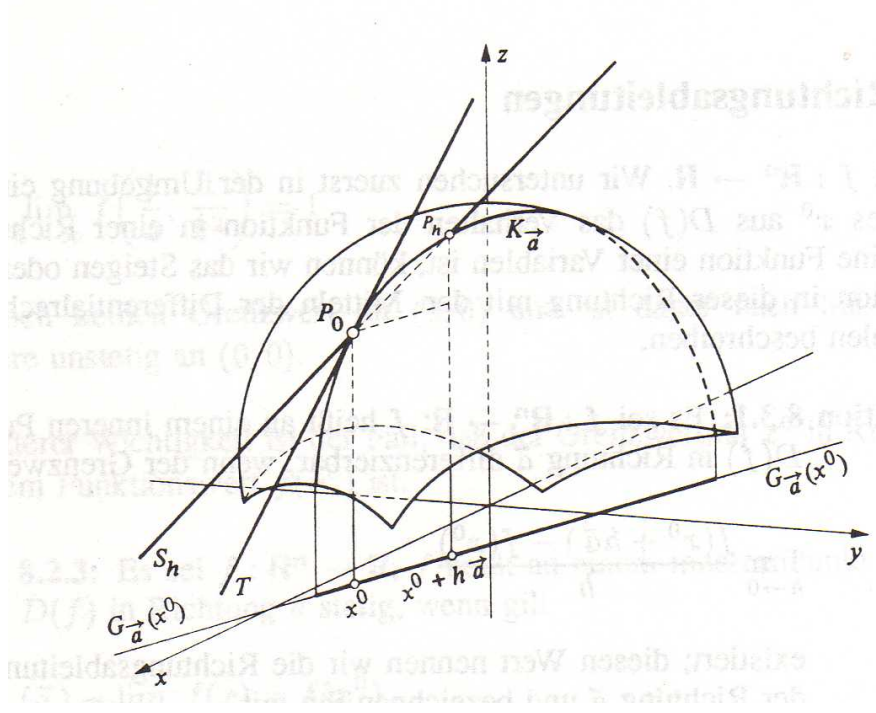


FIGURA 2.3. Ilustración de la derivada direccional.

podemos describir el crecimiento o el decrecimiento de f en esta dirección en términos del cálculo diferencial de una variable.

Definición 2.4. Sea $f : \mathbb{R}^n \rightarrow \mathbb{R}$. La función f se llama diferenciable en un punto interior $x^0 \in D(f)$ en la dirección $\vec{a}$ si el límite

$$\lim_{h \rightarrow 0} \frac{f(x^0 + h\vec{a}) - f(x^0)}{h}$$

existe; si existe, el límite se llama derivada direccional de f en la dirección $\vec{a}$, y lo denotamos por

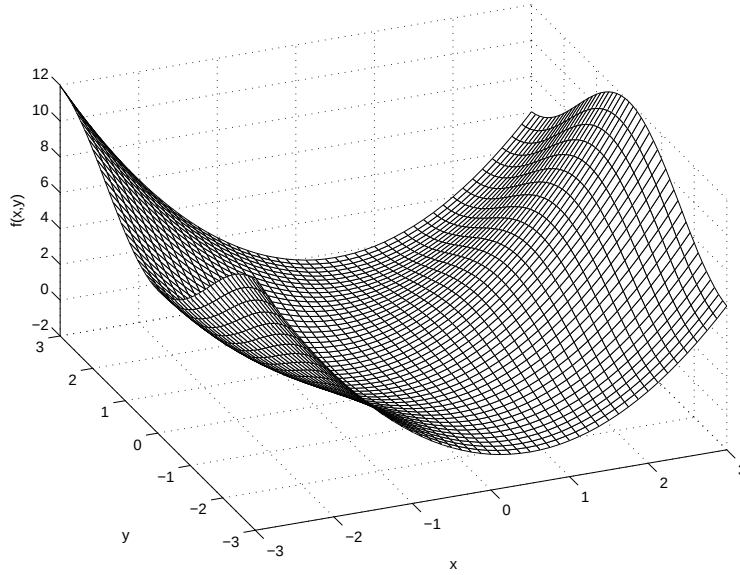
$$\frac{\partial f}{\partial \vec{a}}(x^0).$$

Queremos ilustrar el concepto de la derivada direccional para el caso $n = 2$ (ver Figura 2.3). Sea $x^0 = (x_0, y_0)$ y $\vec{a} = \{a_1, a_2\}$, entonces

$$x^0 + h\vec{a} = (x_0 + ha_1, y_0 + ha_2).$$

En general, el grafo de f es una “superficie” en $\mathbb{R}^3$; pero si consideramos f solamente sobre $G_{\vec{a}}(x^0) \cap D(f)$, el grafo es una “curva” $K_{\vec{a}}$ en esta superficie. Ahora trataremos de encontrar la tangente T a la curva $K_{\vec{a}}$ en el punto $P_0 = (x_0, y_0, f(x_0, y_0))$. Para tal efecto consideramos la secante S_h por los puntos P_0 y

$$P_h = (x_0 + ha_1, y_0 + ha_2, f(x_0 + ha_1, y_0 + ha_2)).$$

FIGURA 2.4. La función $f(x, y) = x^2 + x \cos y$.

Esta secante está únicamente determinada por su pendiente

$$\frac{f(x^0 + h\vec{a}) - f(x^0)}{h}.$$

Ahora, si existe

$$\lim_{h \rightarrow 0} \frac{f(x^0 + h\vec{a}) - f(x^0)}{h} = \frac{\partial f}{\partial \vec{a}}(x^0),$$

entonces la recta por P_0 con la pendiente $\partial f / \partial \vec{a}(x^0)$ —como recta límite de las secantes— es la tangente T a la curva $K_{\vec{a}}$ deseada.

Ejemplo 2.4. Sea la función $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ definida por $f(x, y) = x^2 + x \cos y$ (ver Figura 2.4). Queremos calcular la derivada direccional de f en el punto (x_0, y_0) en la dirección $\vec{a} = (1/\sqrt{2})\{1, 1\}$. Aquí obtenemos

$$\begin{aligned} \frac{\partial f}{\partial \vec{a}}(x_0, y_0) &= \lim_{h \rightarrow 0} \frac{1}{h} \left[\left(x_0 + \frac{h}{\sqrt{2}} \right)^2 + \left(x_0 + \frac{h}{\sqrt{2}} \right) \cos \left(y_0 + \frac{h}{\sqrt{2}} \right) - x_0^2 - x_0 \cos y_0 \right] \\ &= \lim_{h \rightarrow 0} \frac{\sqrt{2}hx_0 + \frac{1}{2}h^2}{h} + \lim_{h \rightarrow 0} \frac{x_0}{\sqrt{2}} \frac{\cos \left(y_0 + \frac{h}{\sqrt{2}} \right) - \cos y_0}{\frac{h}{\sqrt{2}}} + \lim_{h \rightarrow 0} \frac{1}{\sqrt{2}} \cos \left(y_0 + \frac{h}{\sqrt{2}} \right) \\ &= \sqrt{2}x_0 - \frac{x_0 \sin y_0}{\sqrt{2}} + \frac{\cos y_0}{\sqrt{2}}. \end{aligned}$$

Si cambiamos la orientación de la recta, es decir, si consideramos la dirección $-\vec{a}$, se obtiene el siguiente resultado.

Teorema 2.1. *Sea $f : \mathbb{R}^n \rightarrow \mathbb{R}$. Si f es diferenciable en x^0 en la dirección $\vec{a}$, entonces f también es diferenciable en x^0 en la dirección $-\vec{a}$, y*

$$\frac{\partial f}{\partial(-\vec{a})}(x^0) = -\frac{\partial f}{\partial \vec{a}}(x^0).$$

Demostración. Se tiene que

$$\frac{\partial f}{\partial(-\vec{a})}(x^0) = \lim_{h \rightarrow 0} \frac{f(x^0 + h(-\vec{a})) - f(x^0)}{h} = -\lim_{h \rightarrow 0} \frac{f(x^0 + (-h)\vec{a}) - f(x^0)}{-h} = -\frac{\partial f}{\partial \vec{a}}(x^0). \quad \blacksquare$$

Teorema 2.2. *Sea $f : \mathbb{R}^n \rightarrow \mathbb{R}$. Si f es diferenciable en la dirección $\vec{a}$, entonces f es continua en x^0 en la dirección $\vec{a}$.*

Demostración. Se tiene que

$$\lim_{h \rightarrow 0} (f(x^0 + h\vec{a}) - f(x^0)) = \lim_{h \rightarrow 0} \left(\frac{f(x^0 + h\vec{a}) - f(x^0)}{h} \cdot h \right) = \frac{\partial f}{\partial \vec{a}}(x^0) \cdot 0 = 0,$$

lo que implica la afirmación. \blacksquare

2.4. Derivadas parciales

Las derivadas direccionales con respecto a los vectores $\vec{e}_1, \dots, \vec{e}_n$ de la base natural de $\mathbb{V}^n$, las *derivadas parciales*, son muy importantes.

Definición 2.5. *Sea $f : \mathbb{R}^n \rightarrow \mathbb{R}$. La función f se llama parcialmente diferenciable con respecto a x_k en un punto interior x^0 de $D(f)$ si existe la derivada direccional $(\partial f / \partial \vec{e}_k)(x^0)$. También escribimos*

$$\frac{\partial f}{\partial \vec{e}_k}(x^0) = \frac{\partial f}{\partial x_k}(x^0) = f_{x_k}(x^0).$$

Esta expresión se llama derivada parcial (de primer orden) de f con respecto a la variable x_k en el punto x^0 .

Comentamos que según la Definición 2.5,

$$\begin{aligned} f_{x_k}(x^0) &= \lim_{h \rightarrow 0} \frac{f(x^0 + h\vec{e}_k) - f(x^0)}{h} \\ &= \lim_{h \rightarrow 0} \frac{f(x_1^0, \dots, x_{k-1}^0, x_k^0 + h, x_{k+1}^0, \dots, x_n^0) - f(x_1^0, \dots, x_k^0, \dots, x_n^0)}{h}, \end{aligned}$$

es decir, podemos obtener esta derivada parcial fijando $x_1^0, \dots, x_{k-1}^0, x_{k+1}^0, \dots, x_n^0$ y formando la derivada ordinaria con respecto a x_k en x_k^0 .

Ejemplo 2.5.

1. Consideremos sobre $\mathbb{R}^2$ la función $f(x, y) = x^2 + x \cos y$ (ver Figura 2.4). Para $(x_0, y_0) \in \mathbb{R}^2$ se tiene aquí

$$\frac{\partial f}{\partial x}(x_0, y_0) = 2x_0 + \cos y_0, \quad \frac{\partial f}{\partial y}(x_0, y_0) = -x_0 \sin y_0.$$

2. Consideremos nuevamente la función

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad f(x, y) = \begin{cases} \frac{xy}{x^2 + y^2} & \text{si } (x, y) \neq (0, 0), \\ 0 & \text{si } (x, y) = (0, 0). \end{cases}$$

Aquí obtenemos para $(\xi, \eta) \neq (0, 0)$:

$$\frac{\partial f}{\partial x}(\xi, \eta) = \frac{\eta}{\xi^2 + \eta^2} - \frac{2\xi^2\eta}{(\xi^2 + \eta^2)^2}, \quad \frac{\partial f}{\partial y}(\xi, \eta) = \frac{\xi}{\xi^2 + \eta^2} - \frac{2\xi\eta^2}{(\xi^2 + \eta^2)^2},$$

y para $(\xi, \eta) = (0, 0)$:

$$\frac{\partial f}{\partial x}(0, 0) = \lim_{h \rightarrow 0} \frac{\frac{h \cdot 0}{h^2 + 0^2} - 0}{h} = 0, \quad \frac{\partial f}{\partial y}(0, 0) = \lim_{h \rightarrow 0} \frac{\frac{0 \cdot h}{0^2 + h^2} - 0}{h} = 0,$$

es decir, las derivadas parciales existen en cada punto. Por otro lado, ya vimos en el Ejemplo 2.3 que la función no es continua en el punto $(0, 0)$. Sin embargo, note que las derivadas parciales no son acotadas en una vecindad de $(0, 0)$, puesto que para $\xi \neq 0$, $\eta \neq 0$ se tiene que

$$\frac{\partial f}{\partial x}(0, \eta) = \frac{1}{\eta}, \quad \frac{\partial f}{\partial y}(\xi, 0) = \frac{1}{\xi}.$$

Entonces, la existencia de las derivadas parciales aún no implica la continuidad de la función. En el siguiente teorema demostraremos que si todas las derivadas parciales son acotadas, sí podemos concluir que f es continua.

Teorema 2.3. Sea $f : \mathbb{R}^n \rightarrow \mathbb{R}$ y $x^0 \in D(f)$. Si las derivadas parciales $f_{x_1}, \dots, f_{x_n}$ existen en una vecindad $U_r(x^0)$ de x^0 con $U_r(x^0) \subset D(f)$ y son acotadas en esta vecindad, entonces f es continua en x^0 .

Demostración. Sea $\vec{v} = \{v_1, \dots, v_n\} = v_1\vec{e}_1 + \dots + v_n\vec{e}_n$ un vector con $\|\vec{v}\| < r$. Si definimos

$$\vec{v}^0 := \vec{0}; \quad \vec{v}^\nu = \sum_{i=1}^{\nu} v_i \vec{e}_i, \quad \nu = 1, \dots, n,$$

entonces siempre se tiene que

$$\|\vec{v}^\nu\| \leq \|\vec{v}\| < r,$$

y sabemos que

$$x^0 + \vec{v}^\nu \in U_r(x^0), \quad \nu = 0, 1, \dots, n.$$

Ahora consideramos que

$$f(x^0 + \vec{v}) - f(x^0) = (f(x^0 + \vec{v}^n) - f(x^0 + \vec{v}^{n-1})) + (f(x^0 + \vec{v}^{n-1}) - f(x^0 + \vec{v}^{n-2}))$$

$$\begin{aligned}
& + \cdots + (f(x^0 + \vec{v}^1) - f(x^0 + \vec{v}^0)) \\
& = \sum_{\nu=1}^n (f(x^0 + \vec{v}^\nu) - f(x^0 + \vec{v}^{\nu-1})).
\end{aligned}$$

Según el Teorema del Valor Intermedio del cálculo diferencial existen puntos $\xi^\nu \in U_r(x^0)$ localizados en el segmento lineal que une $x^0 + \vec{v}^{\nu-1}$ con $x^0 + \vec{v}^\nu$ tales que

$$f(x^0 + \vec{v}^\nu) - f(x^0 + \vec{v}^{\nu-1}) = v_\nu \cdot f_{x_\nu}(\xi^\nu),$$

luego

$$f(x^0 + \vec{v}) - f(x^0) = \sum_{\nu=1}^n v_\nu \cdot f_{x_\nu}(\xi^\nu).$$

Dado que todas las derivadas parciales son acotadas, existen constantes $M_\nu > 0$ con $|f_{x_\nu}(x)| \leq M_\nu$ para todo $x \in U_r(x^0)$ y $\nu = 1, \dots, n$. Ahora, si $\varepsilon > 0$ elegimos un δ_ε tal que $0 < \delta_\varepsilon < r$ y

$$\delta_\varepsilon < \frac{\varepsilon}{M_1 + \cdots + M_n}.$$

Así, todos los vectores $\vec{v}$ con $\|\vec{v}\| < \delta_\varepsilon$ satisfacen lo siguiente:

$$|f(x^0 + \vec{v}) - f(x^0)| \leq \sum_{\nu=1}^n |v_\nu| M_\nu < \delta_\varepsilon \sum_{\nu=1}^n M_\nu < \varepsilon.$$

Esto implica que f es continua en x^0 . ■

Para cada punto $x^0 \in D(f)$ donde existen las derivadas parciales $f_{x_1}, \dots, f_{x_n}$ existe entonces el vector $\{f_{x_1}(x^0), \dots, f_{x_n}(x^0)\}$.

Definición 2.6. Sea la función $f : \mathbb{R}^n \rightarrow \mathbb{R}$ parcialmente diferenciable con respecto a cada una de las variables x_k , $k = 1, \dots, n$. Entonces el vector

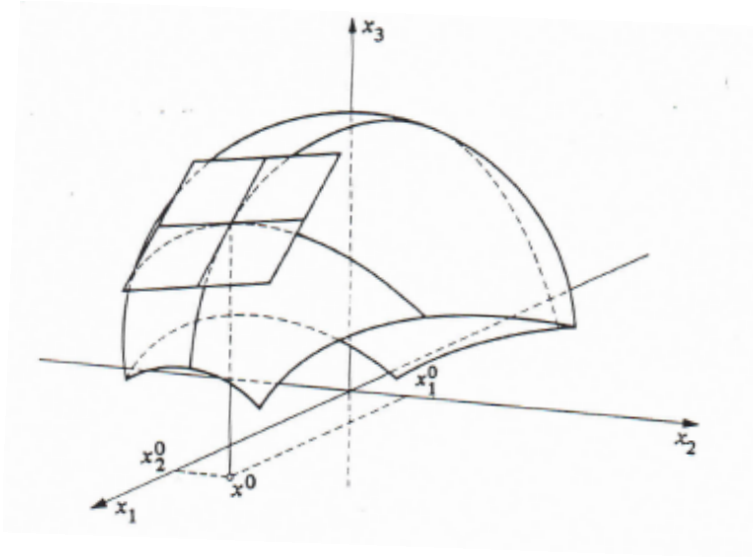
$$\text{grad } f(x^0) = \{f_{x_1}(x^0), f_{x_2}(x^0), \dots, f_{x_n}(x^0)\}$$

se llama gradiente de f en x^0 . Otra notación es

$$\nabla f(x^0) = \text{grad } f(x^0).$$

2.5. La diferenciabilidad en $\mathbb{R}^n$

Los conceptos de la derivada direccional y en particular de la derivada parcial aún no corresponden al concepto de la diferenciabilidad de funciones de una variable dado que la derivada direccional no considera enteramente el comportamiento de la función en una vecindad n -dimensional. Ya sabemos que la existencia de todas las derivadas direccionales en un punto x^0 asegura la continuidad de f en todas las direcciones, pero tal como vimos en el Ejemplo 2.5, esto todavía no nos permite deducir la continuidad de f en el punto x^0 (mientras que para las funciones de una variable, la diferenciabilidad sí implica la continuidad). Ahora introduciremos un concepto de diferenciabilidad que considera enteramente la vecindad n -dimensional.

FIGURA 2.5. La diferenciabilidad de una función $f : \mathbb{R}^2 \rightarrow \mathbb{R}$.

Definición 2.7. Sea $f : \mathbb{R}^n \rightarrow \mathbb{R}$ y sea x^0 un punto interior de $D(f)$. La función se llama diferenciable en el punto x^0 si existen un vector $\vec{c} = \{c_1, \dots, c_n\}$ y una función f^0 definida en una vecindad $U(x^0)$ de x^0 con las siguientes propiedades:

1. $f^0(x^0) = \lim_{x \rightarrow x^0} f^0(x) = 0$.
2. $f(x) = f(x^0) + \vec{c} \cdot (\vec{x} - \vec{x}^0) + d(x, x^0)f^0(x)$ para todo $x \in U(x^0)$.

El vector $\vec{c}$ se llama derivada (o derivada total) de f en x^0 .

Para $n = 1$ este resultado implica que si f es una función diferenciable en x^0 , entonces en una vecindad de x^0 la función f puede ser aproximada por una recta g con $g(x) = f(x^0) + c(x - x^0)$ de tal manera que la diferencia $f(x) - g(x)$ desaparece de por lo menos primer orden cuando $x \rightarrow x^0$, es decir,

$$\frac{f(x) - g(x)}{|x - x^0|} \rightarrow 0 \quad \text{cuando } x \rightarrow x^0.$$

Podemos ofrecer una interpretación análoga para funciones $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ (ver Figura 2.5). Aquí

$$\begin{aligned} g(x) &= f(x^0) + \vec{c} \cdot (\vec{x} - \vec{x}^0) \\ &= f(x_1^0, x_2^0) + c_1(x_1 - x_1^0) + c_2(x_2 - x_2^0) \end{aligned}$$

representa un plano en $\mathbb{R}^3$ por $(x_1^0, x_2^0, f(x_1^0, x_2^0))$. Ahora, la diferenciabilidad en x^0 significa que en una vecindad de x^0 podemos aproximar f por un plano, el *plano tangencial*, de tal manera que la diferencia $f(x) - g(x)$ desaparece de por lo menos primer orden cuando $x \rightarrow x^0$, es decir,

$$\frac{f(x_1, x_2) - g(x_1, x_2)}{d((x_1, x_2), (x_1^0, x_2^0))} \rightarrow 0 \quad \text{cuando } (x_1, x_2) = x \rightarrow x^0 = (x_1^0, x_2^0).$$

Ahora demostraremos que la diferenciabilidad en x^0 implica la existencia de las derivadas parciales $f_{x_1}(x^0), \dots, f_{x_n}(x^0)$.

Teorema 2.4. *Sea $f : \mathbb{R}^n \rightarrow \mathbb{R}$, y sea f diferenciable en x^0 . Entonces todas las derivadas parciales de primer orden existen en x^0 , y*

$$f_{x_k}(x^0) = c_k, \quad k = 1, \dots, n,$$

es decir,

$$\vec{c} = \nabla f(x^0).$$

Demostración. Dado que la función f es diferenciable, se tiene en una vecindad de x^0 que

$$f(x) = f(x^0) + \vec{c} \cdot (\vec{x} - \vec{x}^0) + d(x, x^0)f^0(x).$$

Para un índice k , $1 \leq k \leq n$, sea $x = x^0 + h\vec{e}_k$ con algún $h \neq 0$. Obtenemos $\vec{c} \cdot (\vec{x} - \vec{x}^0) = hc_k$ y $d(x, x^0) = |h|$, por lo tanto

$$\frac{f(x^0 + h\vec{e}_k) - f(x^0)}{h} = c_k + \frac{|h|}{h} f^0(x^0 + h\vec{e}_k).$$

Concluimos que

$$\frac{\partial f}{\partial x_k}(x^0) = \lim_{h \rightarrow 0} \left(c_k + \frac{|h|}{h} f^0(x^0 + h\vec{e}_k) \right) = c_k.$$

■

Más generalmente, la diferenciabilidad implica la existencia de las derivadas direccionales en todas las direcciones. En el teorema siguiente demostraremos, además, como podemos calcular las derivadas direccionales de las derivadas parciales.

Teorema 2.5. *Sea $f : \mathbb{R}^n \rightarrow \mathbb{R}$ diferenciable en x^0 . Entonces f es diferenciable en x^0 en cada dirección $\vec{a}$, y se tiene que*

$$\frac{\partial f}{\partial \vec{a}}(x^0) = (\nabla f(x^0)) \cdot \vec{a}, \tag{2.1}$$

$$\left| \frac{\partial f}{\partial \vec{a}}(x^0) \right| \leq \|\nabla f(x^0)\|. \tag{2.2}$$

Demostración.

1. Según el Teorema 2.4,

$$f(x) = f(x^0) + \vec{c} \cdot (\vec{x} - \vec{x}^0) + d(x, x^0)f^0(x)$$

en una vecindad de x^0 , donde $\vec{c} = \nabla f(x^0)$. Si definimos $x = x^0 + h\vec{a}$ y elegimos $|h| = d(x, x^0) \neq 0$ suficientemente pequeño, se tiene que

$$f(x^0 + h\vec{a}) - f(x^0) = h\vec{c} \cdot \vec{a} + |h|f^0(x^0 + h\vec{a}),$$

por lo tanto

$$\frac{\partial f}{\partial \vec{a}}(x^0) = \lim_{h \rightarrow 0} \frac{f(x^0 + h\vec{a}) - f(x^0)}{h}$$

$$= \lim_{h \rightarrow 0} \left(\vec{c} \cdot \vec{a} + \frac{|h|}{h} f^0(x^0 + h\vec{a}) \right) = \vec{c} \cdot \vec{a} = (\nabla f(x^0)) \cdot \vec{a}.$$

2. Utilizando la desigualdad de Schwarz obtenemos

$$|(\nabla f(x^0)) \cdot \vec{a}| \leq \|\nabla f(x^0)\| \cdot \|\vec{a}\| = \|\nabla f(x^0)\|.$$

■

Comentamos que si $\nabla f(x^0) = \vec{0}$, entonces la fórmula (2.1) implica que

$$\frac{\partial f}{\partial \vec{a}}(x^0) = 0 \quad \text{para toda dirección } \vec{a}.$$

Por otro lado, si $\nabla f(x^0) \neq \vec{0}$, entonces

$$\vec{a}_0 = \frac{1}{\|\nabla f(x^0)\|} \nabla f(x^0)$$

define una dirección en $\mathbb{R}^n$. Para la derivada direccional de f en x^0 en la dirección $\vec{a}_0$ obtenemos

$$\frac{\partial f}{\partial \vec{a}_0}(x^0) = (\nabla f(x^0)) \cdot \vec{a}_0 = \frac{\nabla f(x^0) \cdot \nabla f(x^0)}{\|\nabla f(x^0)\|} = \|\nabla f(x^0)\|.$$

Esto significa que $\vec{a}_0$ es una *dirección extremal*, dado que según la fórmula (2.2),

$$\left| \frac{\partial f}{\partial \vec{a}}(x^0) \right| \leq \|\nabla f(x^0)\| = \frac{\partial f}{\partial \vec{a}_0}(x^0)$$

para *cualquier* dirección $\vec{a}$. En otras palabras, $\vec{a}_0$ es la dirección del mayor crecimiento de f en el punto x^0 .

Si f es diferenciable en un punto x^0 , entonces la existencia de las derivadas direccionales (garantizada por el Teorema 2.5) en x^0 asegura que f es continua en x^0 en todas las direcciones. Sin embargo, esto *no* nos permite concluir que f es continua en x^0 . Pero en el siguiente teorema demostraremos que si f es diferenciable en x^0 (en el sentido de la Definición 2.7), f es continua en x^0 . Para tal efecto demostraremos primero el siguiente teorema.

Teorema 2.6. Sea $f : \mathbb{R}^n \rightarrow \mathbb{R}$, y sea f diferenciable en x^0 . Entonces para cada $\varepsilon > 0$ existe un $\delta > 0$ tal que para todo $x \in D(f)$ con $d(x, x^0) < \delta$ se tiene que

$$|f(x) - f(x^0)| \leq M d(x, x^0)$$

con la constante

$$M = \|\nabla f(x^0)\| + \varepsilon.$$

Demostración. Dado que la función f es diferenciable, existe una vecindad $U(x^0)$ de x^0 donde

$$f(x) = f(x^0) + \vec{c} \cdot (\vec{x} - \vec{x}^0) + d(x, x^0) f^0(x), \quad \vec{c} = \nabla f(x^0).$$

Ahora, para $\varepsilon > 0$ elegimos un $\delta > 0$ tal que $x \in U_\delta(x^0)$ y $|f^0(x)| < \varepsilon$ para todo x tal que $d(x, x^0) < \delta$. Utilizando la desigualdad de Schwarz, obtenemos para $d(x, x^0) < \delta$ lo siguiente:

$$\begin{aligned} |f(x) - f(x^0)| &\leq |\vec{c} \cdot (\vec{x} - \vec{x}^0)| + d(x, x^0) |f^0(x)| \\ &\leq \|\vec{c}\| \|\vec{x} - \vec{x}^0\| + d(x, x^0) |f^0(x)| \end{aligned}$$

$$\leq (\|\nabla f(x^0)\| + \varepsilon)d(x, x^0).$$

■

Con la ayuda del Teorema 2.6 podemos demostrar ahora la continuidad de f .

Teorema 2.7. *Sea $f : \mathbb{R}^n \rightarrow \mathbb{R}$ diferenciable en x^0 . Entonces f es continua en x^0 .*

Demostración. Sea $\{x^k\}_{k \in \mathbb{N}}$ una sucesión en $D(f)$ con $x^k \rightarrow x^0$ cuando $k \rightarrow \infty$. Según el Teorema 2.6 existe una constante M tal que para cada k suficientemente grande

$$|f(x^k) - f(x^0)| \leq Md(x^k, x^0).$$

Esto implica que $f(x^k) \rightarrow f(x^0)$. ■

Si todas las derivadas parciales de una función f existen en un punto x^0 , no necesariamente f debe ser diferenciable en x^0 , dado que en este caso ni siquiera podemos concluir que f es continua en x^0 . Pero demostraremos bajo una hipótesis adicional la siguiente inversión del Teorema 2.4.

Teorema 2.8. *Sea $f : \mathbb{R}^n \rightarrow \mathbb{R}$. Si en una vecindad $U(x^0)$ de x^0 las derivadas parciales $f_{x_1}, \dots, f_{x_n}$ existen y son continuas en x^0 , entonces f es diferenciable en x^0 .*

Demostración. Existe un $\varepsilon_0 > 0$ tal que $U_{\varepsilon_0}(x^0) \subset U(x^0) \subset D(f)$. Elegimos $x = x^0 + \vec{v}$ con $0 \leq \|\vec{v}\| < \varepsilon_0$, donde

$$\vec{v} = \{v_1, \dots, v_n\} = \sum_{i=1}^n v_i \vec{e}_i;$$

además definimos

$$\vec{v}^0 := \vec{0}; \quad \vec{v}^\nu = \sum_{i=1}^\nu v_i \vec{e}_i, \quad \nu = 1, \dots, n,$$

entonces siempre se tiene que

$$\|\vec{v}^\nu\| \leq \|\vec{v}\| < \varepsilon_0,$$

y sabemos que

$$x^0 + \vec{v}^\nu \in U_{\varepsilon_0}(x^0), \quad \nu = 0, 1, \dots, n.$$

Tal como en la demostración del Teorema 2.3, existen ciertos puntos ξ^ν en el segmento lineal que une $x^0 + \vec{v}^{\nu-1}$ con $x^0 + \vec{v}^\nu$ tales que

$$\begin{aligned} f(x^0 + \vec{v}) &= f(x^0) + \sum_{\nu=1}^n (f(x^0 + \vec{v}^\nu) - f(x^0 + \vec{v}^{\nu-1})) \\ &= f(x^0) + \sum_{\nu=1}^n v_\nu \frac{\partial f}{\partial x_\nu}(x^0) + \sum_{\nu=1}^n v_\nu \left(\frac{\partial f}{\partial x_\nu}(\xi^\nu) - \frac{\partial f}{\partial x_\nu}(x^0) \right) \\ &= f(x^0) + \sum_{\nu=1}^n v_\nu \frac{\partial f}{\partial x_\nu}(x^0) + \|\vec{v}\| \sum_{\nu=1}^n \frac{v_\nu}{\|\vec{v}\|} \left(\frac{\partial f}{\partial x_\nu}(\xi^\nu) - \frac{\partial f}{\partial x_\nu}(x^0) \right). \end{aligned}$$

Para demostrar el teorema solamente hay que demostrar que la siguiente función es continua en x^0 :

$$f^0(x) = f^0(x^0 + \vec{v}) = \begin{cases} \sum_{\nu=1}^n \frac{v_\nu}{\|\vec{v}\|} \left(\frac{\partial f}{\partial x_\nu}(\xi^\nu) - \frac{\partial f}{\partial x_\nu}(x^0) \right) & \text{si } \vec{v} \neq \vec{0}, \\ 0 & \text{si } \vec{v} = \vec{0}. \end{cases}$$

Dado que las derivadas parciales $\partial f / \partial x_\nu$ son continuas en x^0 , para cada $\varepsilon > 0$ existe un $\delta_\varepsilon(\nu)$ con $0 < \delta_\varepsilon(\nu) < \varepsilon_0$ tal que

$$\left| \frac{\partial f}{\partial x_\nu}(x) - \frac{\partial f}{\partial x_\nu}(x^0) \right| < \frac{\varepsilon}{n} \quad \text{para todo } x \text{ con } d(x, x^0) < \delta_\varepsilon(\nu).$$

Ahora elegimos

$$\delta_\varepsilon = \min\{\delta_\varepsilon(1), \dots, \delta_\varepsilon(n)\}.$$

Así, para todo x con $d(x, x^0) < \delta_\varepsilon$ también se tiene

$$d(\xi^\nu, x^0) \leq \|\vec{v}\| < \delta_\varepsilon,$$

y finalmente llegamos a

$$|f^0(x)| \leq \sum_{\nu=1}^n \frac{|v_\nu|}{\|\vec{v}\|} \left| \frac{\partial f}{\partial x_\nu}(\xi^\nu) - \frac{\partial f}{\partial x_\nu}(x^0) \right| \leq \sum_{\nu=1}^n \left| \frac{\partial f}{\partial x_\nu}(\xi^\nu) - \frac{\partial f}{\partial x_\nu}(x^0) \right| < \varepsilon,$$

es decir, f^0 es continua en x^0 . ■

Ejemplo 2.6. Consideremos nuevamente la función $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ dada por $f(x, y) = x^2 + x \cos y$. Las derivadas parciales

$$\frac{\partial f}{\partial x}(x, y) = 2x + \cos y, \quad \frac{\partial f}{\partial y}(x, y) = -x \sin y$$

son continuas en cada punto $(x_0, y_0) \in \mathbb{R}^2$, por lo tanto f es diferenciable en todo $\mathbb{R}^2$. Ahora, si nuevamente nos ponemos la tarea de calcular la derivada direccional de f en (x_0, y_0) en la dirección $\vec{a} = (1/\sqrt{2})\{1, 1\}$, obtenemos aplicando el resultado del Teorema 2.5

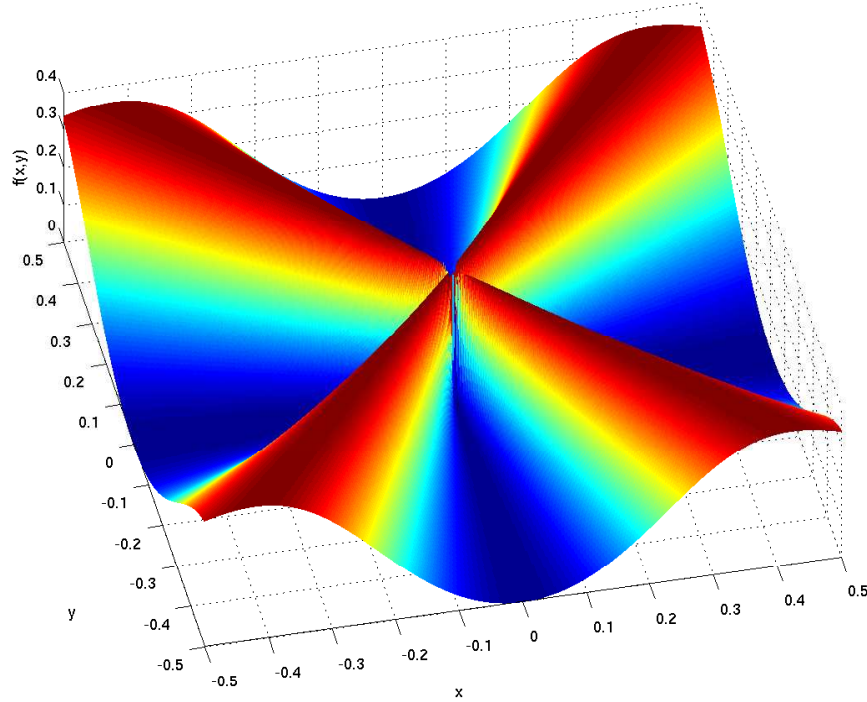
$$\begin{aligned} \frac{\partial f}{\partial \vec{a}}(x_0, y_0) &= (\nabla f(x_0, y_0)) \cdot \vec{a} = \{2x_0 + \cos y_0, -x_0 \sin y_0\} \cdot \left(\frac{1}{\sqrt{2}}\{1, 1\} \right) \\ &= \sqrt{2}x_0 + \frac{\cos y_0}{\sqrt{2}} - \frac{x_0 \sin y_0}{\sqrt{2}}, \end{aligned}$$

lo que reconfirma el resultado del Ejemplo 2.4 por una computación mucho más compacta.

Ejemplo 2.7 (Problema 1, Certamen 1, Curso 2009/I). Se considera la función $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ dada por

$$f(x, y) = \begin{cases} \frac{x^2 y^2}{2x^4 + y^4} & \text{si } (x, y) \neq (0, 0), \\ 0 & \text{si } (x, y) = (0, 0). \end{cases}$$

a) Determinar las derivadas parciales de f .

FIGURA 2.6. Función $f(x, y)$ del Ejemplo 2.7.

- b) ¿La función f es diferenciable en $(0, 0)$?
- c) Sea $(x_0, y_0) = (1, 1)$. Determinar la derivada direccional de f en (x_0, y_0) en la dirección $\vec{a} = (0, 6, 0, 8)$.
- d) Determinar en (x_0, y_0) la dirección de mayor crecimiento de f .

Solución sugerida. Se incluye un plot de la función f , ver Figura 2.6.

- a) Sea $(x, y) \neq 0$, entonces podemos calcular las derivadas parciales directamente:

$$\frac{\partial f}{\partial x}(x, y) = \frac{2xy^2(2x^4 + y^4) - x^2y^2 \cdot 8x^3}{(2x^4 + y^4)^2} = 2 \frac{xy^2}{2x^4 + y^4} - \frac{8x^5y^2}{(2x^4 + y^4)^2},$$

$$\frac{\partial f}{\partial y}(x, y) = \frac{2x^2y(2x^4 + y^4) - x^2y^2 \cdot 4y^3}{(2x^4 + y^4)^2} = 2 \frac{x^2y}{2x^4 + y^4} - \frac{4x^2y^5}{(2x^4 + y^4)^2}.$$

En virtud de

$$\lim_{h \rightarrow 0} \frac{f(h, 0) - f(0, 0)}{h} = \lim_{h \rightarrow 0} \frac{0}{2h^4 + 0} = 0,$$

$$\lim_{h \rightarrow 0} \frac{f(0, h) - f(0, 0)}{h} = \lim_{h \rightarrow 0} \frac{0}{0 + h^4} = 0$$

podemos concluir que

$$\frac{\partial f}{\partial x}(0, 0) = 0, \quad \frac{\partial f}{\partial y}(0, 0) = 0.$$

- b) La función f no es diferenciable en $(0,0)$. Para demostrar esto basta probar que f no es continua en $(0,0)$, puesto que según el Teorema 2.7, si f no es continua en un punto (x^0, y^0) , no es diferenciable en este punto. Los siguientes cálculos son suficientes para demostrar que f no es continua en $(0,0)$:

$$\lim_{h \rightarrow 0} f(h, 0) = 0, \quad \lim_{h \rightarrow 0} f(h, h) = \lim_{h \rightarrow 0} \frac{h^4}{2h^4 + h^4} = \lim_{h \rightarrow 0} \frac{1}{3} = \frac{1}{3}.$$

- c) La derivada direccional solicitada es

$$\begin{aligned} \frac{\partial f}{\partial \vec{a}}(x_0, y_0) &= 0,6 \frac{\partial f}{\partial x}(1, 1) + 0,8 \frac{\partial f}{\partial y}(1, 1) \\ &= 0,6 \cdot \left(\frac{2}{3} - \frac{8}{9} \right) + 0,8 \cdot \left(\frac{2}{3} - \frac{4}{9} \right) = \frac{2}{45} = 0,0\bar{4}. \end{aligned}$$

- d) La dirección de mayor crecimiento es

$$\frac{\nabla f(1, 1)}{\|\nabla f(1, 1)\|_2} = \frac{9}{\sqrt{8}} \left(-\frac{2}{9}, \frac{2}{9} \right)^T = \left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} \right)^T.$$

2.6. Derivadas parciales de orden mayor

Sea $f : \mathbb{R}^n \rightarrow \mathbb{R}$. Supongamos que para un índice k ($1 \leq k \leq n$) la derivada parcial f_{x_k} existe sobre un dominio $D(f_{x_k}) \subset D(f)$. En este caso, podemos tratar de formar en un punto interior $x^0 \in D(f_{x_k})$ para un índice l ($1 \leq l \leq n$) la derivada parcial

$$(f_{x_k})_{x_l} = f_{x_k x_l}.$$

Ahora, si a su vez $f_{x_k x_l}$ existe sobre $D(f_{x_k x_l}) \subset D(f_{x_k})$, podemos tratar de formar la derivada parcial

$$(f_{x_k x_l})_{x_m} = f_{x_k x_l x_m} \quad (1 \leq m \leq n),$$

etc. Esta consideración nos lleva a la siguiente definición.

Definición 2.8. Sea $f : \mathbb{R}^n \rightarrow \mathbb{R}$. Si en un punto interior

$$x^0 \in D(f_{x_{k_1} x_{k_2} \dots x_{k_{l-1}}})$$

existe para un k_l con $l > 1$ la derivada parcial

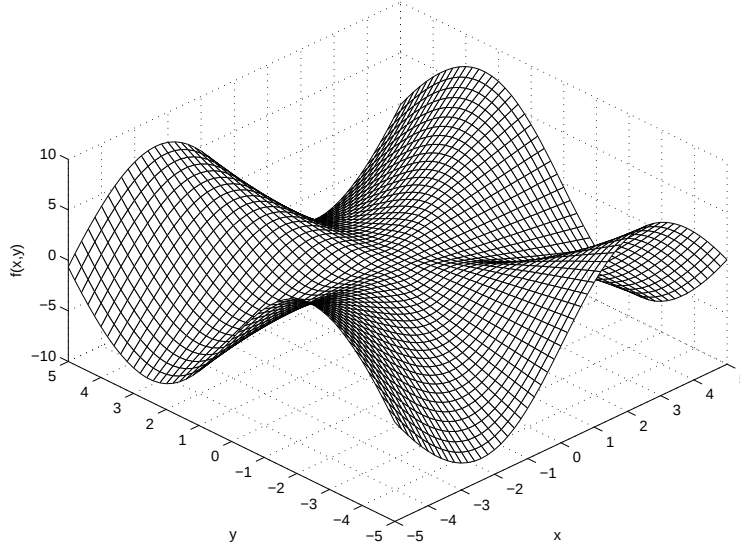
$$(f_{x_{k_1} x_{k_2} \dots x_{k_{l-1}}})_{x_{k_l}}(x^0) \quad (1 \leq k_i \leq n \text{ para } i = 1, \dots, l),$$

entonces esta derivada parcial se llama derivada parcial del orden l de f en el punto x^0 . También usamos la notación

$$f_{x_{k_1} x_{k_2} \dots x_{k_l}}(x^0) \quad \text{o} \quad \frac{\partial^l f}{\partial x_{k_1} \dots \partial x_{k_l}}(x^0).$$

Las derivadas parciales de la Definición 2.5 son derivadas parciales de primer orden; a veces la misma función f se llama derivada parcial del orden cero.

Definición 2.9. Sea $f : \mathbb{R}^n \rightarrow \mathbb{R}$ y $X \subset \mathbb{R}^n$ un conjunto abierto. Sea $k \geq 0$ un número entero. Escribimos $f \in C^k(X)$ si sobre X todas las derivadas de f del orden k existen y son continuas.

FIGURA 2.7. La función $f(x, y)$ del Ejemplo 2.9.

Ejemplo 2.8. Consideremos sobre $\mathbb{R}^3$ la función

$$f(x, y, z) = 4xyz - x^2 + y^2.$$

Aquí obtenemos las derivadas parciales

$$\begin{aligned} f_x &= 4yz - 2x, & f_{xx} &= -2, & f_{xy} &= f_{yx} = 4z, \\ f_y &= 4xz + 2y, & f_{yy} &= 2, & f_{xz} &= f_{zx} = 4y, \\ f_z &= 4xy, & f_{zz} &= 0, & f_{yz} &= f_{zy} = 4x. \end{aligned}$$

En este ejemplo obtenemos $f_{xy} = f_{yx}$, $f_{xz} = f_{zx}$ y $f_{yz} = f_{zy}$. La pregunta es si siempre podemos intercambiar el orden de las derivaciones parciales. Pero esto no es válido en general, como muestra el siguiente ejemplo.

Ejemplo 2.9. Consideremos sobre $\mathbb{R}^2$ la función

$$f(x, y) = \begin{cases} xy \frac{x^2 - y^2}{x^2 + y^2} & \text{si } (x, y) \neq (0, 0), \\ 0 & \text{si } (x, y) = (0, 0), \end{cases}$$

ver Figura 2.7. Aquí obtenemos las derivadas parciales en $(0, 0)$

$$f_x(0, 0) = \lim_{h \rightarrow 0} \frac{f(h, 0) - 0}{h} = 0, \quad f_y(0, 0) = \lim_{h \rightarrow 0} \frac{f(0, h) - 0}{h} = 0,$$

y para $(x, y) \neq (0, 0)$

$$f_x(x, y) = y \cdot \frac{x^4 + 4x^2y^2 - y^4}{(x^2 + y^2)^2}, \quad f_y(x, y) = x \cdot \frac{x^4 - 4x^2y^2 - y^4}{(x^2 + y^2)^2},$$

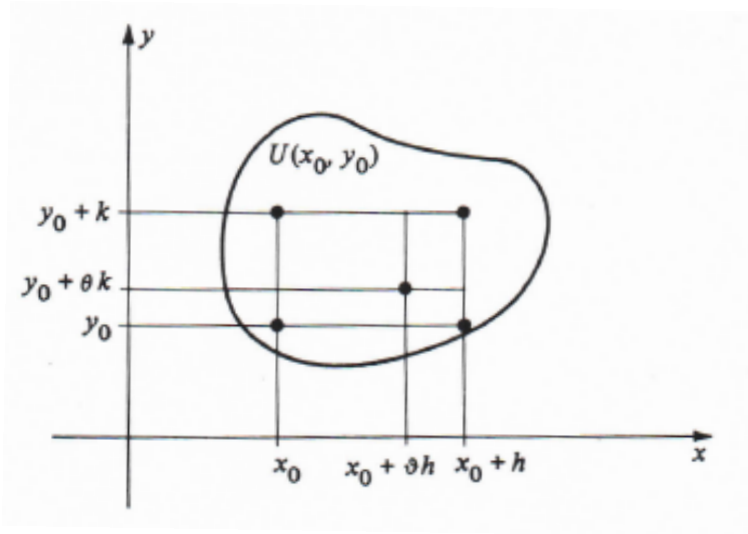


FIGURA 2.8. Ilustración de la demostración del Teorema 2.9.

luego

$$f_{xy}(0,0) = \lim_{h \rightarrow 0} \frac{1}{h} \left[h \cdot \frac{-h^4}{(h^2)^2} \right] = -1, \quad f_{yx}(0,0) = \lim_{h \rightarrow 0} \frac{1}{h} \left[h \cdot \frac{h^4}{(h^2)^2} \right] = 1.$$

Observamos que las segundas derivadas parciales $f_{xy}(0,0)$ y $f_{yx}(0,0)$ existen, pero sus valores son diferentes. En este caso, ambas funciones f_{xy} y f_{yx} son discontinuas en $(0,0)$. Para ver eso, calculamos primero para $(x,y) \neq (0,0)$

$$f_{xy}(x,y) = f_{yx}(x,y) = \frac{(x^2 - y^2)(x^4 + 10x^2y^2 + y^4)}{(x^2 + y^2)^3}.$$

Para la sucesión

$$\{(x_k, y_k)\}_{k \in \mathbb{N}} = \left\{ \left(\frac{1}{k}, \frac{1}{k} \right) \right\}_{k \in \mathbb{N}}$$

obtenemos

$$f_{xy}(x_k, y_k) = f_{yx}(x_k, y_k) = 0,$$

es decir

$$\lim_{k \rightarrow \infty} f_{xy}(x_k, y_k) = 0 \neq f_{xy}(0,0) = -1, \quad \lim_{k \rightarrow \infty} f_{yx}(x_k, y_k) = 0 \neq f_{yx}(0,0) = 1.$$

En lo siguiente queremos analizar bajo qué condiciones podemos intercambiar el orden de las derivaciones parciales. Para tal efecto demostraremos primeramente el siguiente teorema.

Teorema 2.9 (Schwarz). Sea $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ y sea $U(x_0, y_0)$ una vecindad abierta. Supongamos que la derivada parcial f_{xy} existe en $U(x_0, y_0)$ y es continua en (x_0, y_0) ; además supongamos que $f_y(x, y_0)$ existe para todo $(x, y_0) \in U(x_0, y_0)$. Entonces también existe $f_{yx}(x_0, y_0)$, y se tiene que

$$f_{xy}(x_0, y_0) = f_{yx}(x_0, y_0).$$

Demostración. Tenemos que demostrar que la función

$$F(h) := \frac{1}{h} (f_y(x_0 + h, y_0) - f_y(x_0, y_0))$$

satisface

$$\lim_{h \rightarrow 0} F(h) = f_{xy}(x_0, y_0).$$

Si elegimos $h \neq 0$ tal que $(x_0 + h, y_0) \in U(x_0, y_0)$ (ver Figura 2.8), entonces se tiene que

$$\begin{aligned} f_y(x_0 + h, y_0) &= \lim_{k \rightarrow 0} \frac{f(x_0 + h, y_0 + k) - f(x_0 + h, y_0)}{k}, \\ f_y(x_0, y_0) &= \lim_{k \rightarrow 0} \frac{f(x_0, y_0 + k) - f(x_0, y_0)}{k}. \end{aligned}$$

Utilizando la abreviatura

$$G(h, k) := \frac{f(x_0 + h, y_0 + k) - f(x_0 + h, y_0) - [f(x_0, y_0 + k) - f(x_0, y_0)]}{hk}$$

notamos que

$$F(h) = \lim_{k \rightarrow 0} G(h, k).$$

Ahora sea $k \neq 0$ fijo y $(x_0, y_0 + k) \in U(x_0, y_0)$, y sea

$$\varphi(x) := f(x, y_0 + k) - f(x, y_0).$$

De la existencia de f_{xy} sigue la existencia de f_x , por lo tanto según el Teorema del Valor Intermedio del cálculo diferencial existe un número $\vartheta \in (0, 1)$ tal que

$$\varphi(x + h) - \varphi(x) = h\varphi'(x_0 + \vartheta h),$$

es decir,

$$G(h, k) = \frac{f_x(x_0 + \vartheta h, y_0 + k) - f_x(x_0 + \vartheta h, y_0)}{k}.$$

Además sea

$$\psi(y) := f_x(x_0 + \vartheta h, y).$$

La existencia de f_{xy} en $U(x_0, y_0)$ implica la existencia de ψ' , y existe un número $\theta = \theta(k) \in (0, 1)$ (ver Figura 2.8) tal que

$$\psi(y_0 + k) - \psi(y_0) = k\psi'(y_0 + \theta k) = kf_{xy}(x_0 + \vartheta h, y_0 + \theta k),$$

es decir

$$G(h, k) = f_{xy}(x_0 + \vartheta h, y_0 + \theta k).$$

Puesto que f_{xy} es continua en (x_0, y_0) , existe para $\varepsilon > 0$ un $\delta_\varepsilon > 0$ tal que

$$|f_{xy}(x, y) - f_{xy}(x_0, y_0)| < \frac{\varepsilon}{2}$$

para todo (x, y) con $d((x_0, y_0), (x, y)) < \delta$. Esto implica que para todo h y k con $|h|$ y $|k|$ suficientemente pequeño

$$|G(h, k) - f_{xy}(x_0, y_0)| < \frac{\varepsilon}{2},$$

y obtenemos

$$|F(h) - f_{xy}(x_0, y_0)| = \lim_{k \rightarrow 0} |G(h, k) - f_{xy}(x_0, y_0)| \leq \frac{\varepsilon}{2} < \varepsilon.$$

Pero esto significa que

$$\lim_{h \rightarrow 0} F(h) = f_{xy}(x_0, y_0).$$

■

En particular, las hipótesis del Teorema 2.9 están satisfechas si $f \in C^2(U(x_0, y_0))$ para una vecindad abierta $U(x_0, y_0)$ de (x_0, y_0) .

Si ambas derivadas parciales $f_{xy}(x_0, y_0)$ y $f_{yx}(x_0, y_0)$ existen, estos dos valores pueden ser diferentes, según el Teorema 2.9, solamente si ambas funciones f_{xy} y f_{yx} son discontinuas en (x_0, y_0) . Esto sucede en el Ejemplo 2.9.

Teorema 2.10. Sea $f : \mathbb{R}^n \rightarrow \mathbb{R}$ y sea $X \subset \mathbb{R}^n$ un conjunto abierto. Para un $k \geq 1$ sea $f \in C^k(X)$; además sea $\nu_i \in \{1, \dots, k\}$ para $1 \leq i \leq k$. Entonces para cada permutación $\mu_1, \dots, \mu_k$ de los números $\nu_1, \dots, \nu_k$ y todo $x^0 \in X$ se tiene que

$$f_{x_{\nu_1} x_{\nu_2} \dots x_{\nu_k}}(x^0) = f_{x_{\mu_1} x_{\mu_2} \dots x_{\mu_k}}(x^0).$$

Demostración. Dado que cada permutación $\nu_1, \dots, \nu_k$ puede ser generada por un número finito de permutaciones de solamente dos elementos consecutivos, el enunciado del teorema es una consecuencia del Teorema 2.9. ■

Ejemplo 2.10 (Problema 2, Certamen 1, Curso 2009/I). Sea $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ dada por

$$f(x, y) = \begin{cases} xy \frac{x^2 - 2y^2}{4x^2 + y^2} & \text{si } (x, y) \neq (0, 0), \\ 0 & \text{si } (x, y) = (0, 0). \end{cases}$$

- ¿Se puede aplicar el Teorema de Schwarz para f en el punto $(x_0 = 0, y_0 = 0)$?
- Sea el conjunto $\mathcal{E}$ dado por

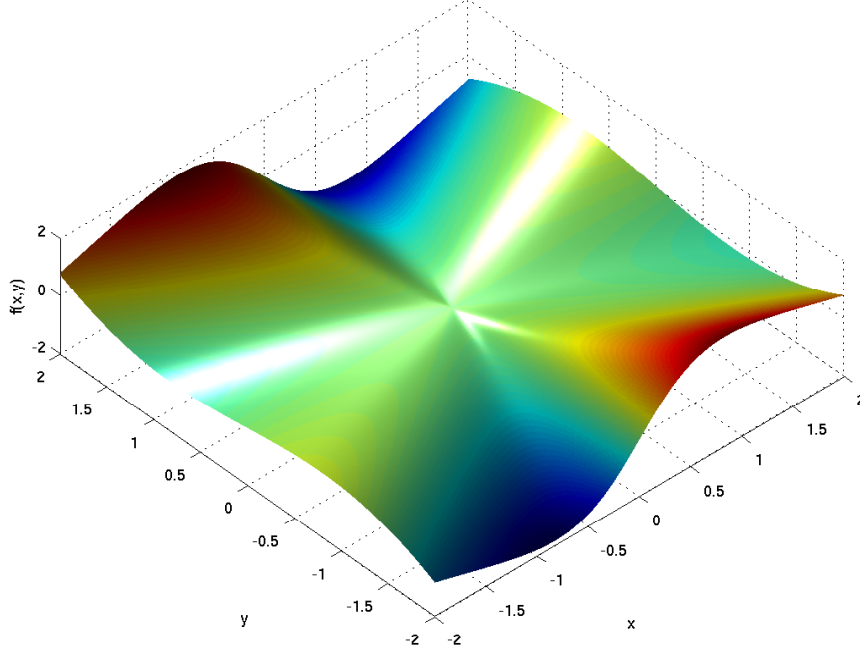
$$\mathcal{E} = \left\{ (x, y) \in \mathbb{R}^2 \mid x^2 + \left(\frac{y}{2}\right)^2 \leq 1 \right\}.$$

¿Cuales son los extremos absolutos de f sobre $\mathcal{E}$?

Solución sugerida. Se incluye un plot de la función f , ver Figura 2.9.

- Para resolver ambas partes del problema se requiere calcular las derivadas parciales. Sea $(x, y) \neq (0, 0)$. Aquí obtenemos

$$\frac{\partial f}{\partial x}(x, y) = y \frac{x^2 - 2y^2}{4x^2 + y^2} + xy \frac{2x(4x^2 + y^2) - (x^2 - 2y^2) \cdot 8x}{(4x^2 + y^2)^2}$$

FIGURA 2.9. Función $f(x, y)$ del Ejemplo 2.10.

$$\begin{aligned}
 &= y \frac{x^2 - 2y^2}{4x^2 + y^2} + x^2 y \frac{8x^2 + 2y^2 - 8x^2 + 16y^2}{(4x^2 + y^2)^2} = y \frac{x^2 - 2y^2}{4x^2 + y^2} + \frac{18x^2 y^3}{(4x^2 + y^2)^2} \\
 &= \frac{(4x^2 y + y^3)(x^2 - 2y^2) + 18x^2 y^3}{(4x^2 + y^2)^2} = y \frac{4x^4 + 11x^2 y^2 - 2y^4}{(4x^2 + y^2)^2}, \\
 \frac{\partial f}{\partial y}(x, y) &= x \frac{x^2 - 2y^2}{4x^2 + y^2} - \frac{18x^3 y^2}{(4x^2 + y^2)^2} = \frac{(4x^3 + xy^2)(x^2 - 2y^2) - 18x^3 y^2}{(4x^2 + y^2)^2} \\
 &= x \frac{4x^4 - 2y^4 - 25x^2 y^2}{(4x^2 + y^2)^2}.
 \end{aligned}$$

Por otro lado, para $(x, y) = 0$ obtenemos

$$\lim_{h \rightarrow 0} \frac{f(h, 0) - f(0, 0)}{h} = 0, \quad \lim_{h \rightarrow 0} \frac{f(0, h) - f(0, 0)}{h} = 0.$$

Entonces, concluimos que

$$\frac{\partial f}{\partial x}(0, 0) = 0, \quad \frac{\partial f}{\partial y}(0, 0) = 0,$$

es decir las primeras derivadas existen en $(0, 0)$. Luego calculamos para $(x, y) \neq (0, 0)$

$$\frac{\partial^2 f}{\partial x \partial y}(x, y) = \frac{4x^4 + 11x^2 y^2 - 2y^4}{(4x^2 + y^2)^2}$$

$$+ y \frac{(-8y^3 + 22x^2)(4x^2 + y^2) - 4y(4x^4 + 11x^2y^2 - 2y^4)}{(4x^2 + y^2)^3},$$

lo que significa que

$$\lim_{x \rightarrow 0} \frac{\partial^2 f}{\partial x \partial y}(x, 0) = \lim_{x \rightarrow 0} \frac{4x^4}{16x^4} = \frac{1}{4},$$

por otro lado

$$\frac{\partial^2 f}{\partial x \partial y}(0, 0) = \lim_{h \rightarrow 0} \frac{1}{h} \left(\frac{\partial f}{\partial x}(0, h) - \frac{\partial f}{\partial x}(0, 0) \right) = \lim_{h \rightarrow 0} \frac{1}{h} \left(h \cdot \frac{-2h^4}{h^4} \right) = -2 \neq \frac{1}{4}.$$

Concluimos que f_{xy} no es continua en $(0, 0)$, por lo tanto el Teorema de Schwarz no puede ser aplicado.

- b) Tratemos primeramente identificar los extremos de f en el interior de $\mathcal{E}$, utilizando la relación

$$\nabla f(x_0, y_0) = 0 \quad (2.3)$$

para identificar los puntos críticos (candidatos a extremos). Para tal efecto, notamos que la ecuación

$$\frac{\partial f}{\partial x}(x, y) = 0,$$

donde la derivada parcial fue calculada arriba, entrega las soluciones

$$y_0 = 0 \quad (2.4)$$

o para $x = x_0$

$$4x^4 + 11x^2y^2 - 2y^4 = 0 \quad (2.5)$$

$$\Leftrightarrow x^4 + \frac{11}{4}x^2y^2 + \frac{121}{64}y^4 = \frac{1}{2}y^4 + \frac{121}{64}y^4 = \frac{153}{64}y^4$$

$$\Leftrightarrow \left(x^2 + \frac{11}{8}y^2 \right)^2 = \frac{153}{64}y^4$$

$$\Leftrightarrow x^2 + \frac{11}{8}y^2 = -\frac{\sqrt{153}}{8}y^2 \quad \vee \quad x^2 + \frac{11}{8}y^2 = \frac{\sqrt{153}}{8}y^2. \quad (2.6)$$

La primera de las dos posibilidades no puede ser satisfecha por ningún $(x_0, y_0) \in \mathbb{R}^2$. Supongamos que (x_0, y_0) satisface la segunda ecuación, entonces

$$x_0^2 = \frac{\sqrt{153} - 11}{8}y_0^2. \quad (2.7)$$

Repitiendo estos mismos cálculos para

$$\frac{\partial f}{\partial y}(x, y) = 0,$$

obtenemos

$$x_0 = 0 \quad (2.8)$$

o

$$x_0^2 = \frac{\sqrt{593} + 25}{8} y_0^2 \quad \vee \quad x_0^2 = \frac{-\sqrt{593} + 25}{8} y_0^2. \quad (2.9)$$

Así hemos demostrado que cualquier solución de (2.3) debe satisfacer una de las relaciones (2.4) y (2.7) y una de las relaciones (2.8) y (2.9). El único punto crítico que obtenemos así es $(x_0 = 0, y_0 = 0)$, pero este punto no es un máximo ni un mínimo, puesto que en cada vecindad de $(0, 0)$ existen valores de f positivos tanto que negativos. Para ver eso, sea $\varepsilon > 0$, entonces

$$f(\varepsilon, \varepsilon) = -\frac{\varepsilon^2}{5} < 0, \quad f(2\varepsilon, \varepsilon) = \frac{4}{17}\varepsilon^2 > 0.$$

Analicemos ahora el comportamiento de f sobre la frontera $\partial\mathcal{E}$. Para tal efecto, notamos que $(x, y) \in \partial\mathcal{E}$ corresponde a la satisfacción de la restricción

$$g(x, y) := 4x^2 + y^2 - 4 = 0. \quad (2.10)$$

Tomando en cuenta que sobre $\mathcal{E}$ tenemos que

$$\begin{aligned} \frac{\partial f}{\partial x}(x, y) &= y \left(\frac{x^2 - 2y^2}{4} + \frac{9x^2 y^2}{8} \right), \\ \frac{\partial f}{\partial y}(x, y) &= x \left(\frac{x^2 - 2y^2}{4} - \frac{9x^2 y^2}{8} \right) \quad \text{para } (x, y) \in \partial\mathcal{E}. \end{aligned}$$

Las ecuaciones $\nabla f(x, y) - \lambda \nabla g(x, y) = 0$, donde λ es el multiplicador de Lagrange, son

$$y \left(\frac{x^2 - 2y^2}{4} + \frac{9x^2 y^2}{8} \right) - 8\lambda x = 0, \quad (2.11)$$

$$x \left(\frac{x^2 - 2y^2}{4} - \frac{9x^2 y^2}{8} \right) - 2\lambda y = 0. \quad (2.12)$$

Multiplicando (2.11) por $16y$, (2.12) por $(-64x)$, sumando las dos ecuaciones y tomando en cuenta que $y^2 + 4x^2 = 4$ obtenemos

$$(y^2 - 4x^2)(x^2 - 2y^2) + 18x^2 y^2 = 0.$$

Substituyendo $y^2 = 4 - 4x^2$ y luego poniendo $z := x^2$ obtenemos

$$(1 - 2z)(9z - 8) + 18z(1 - z) = 0,$$

es decir

$$-36z^2 + 43z = 8.$$

Las soluciones de esta ecuación cuadrática son

$$z_1 = \frac{43 + \sqrt{697}}{72}, \quad z_2 = \frac{43 - \sqrt{697}}{72},$$

las que dan origen a los siguientes ocho puntos críticos localizados en $\partial\mathcal{E}$:

$$\begin{aligned} P_1 &= (\sqrt{z_1}, \sqrt{4-4z_1}), & P_5 &= (-\sqrt{z_1}, \sqrt{4-4z_1}), \\ P_2 &= (\sqrt{z_2}, \sqrt{4-4z_2}), & P_6 &= (-\sqrt{z_2}, \sqrt{4-4z_2}), \\ P_3 &= (\sqrt{z_1}, -\sqrt{4-4z_1}), & P_7 &= (-\sqrt{z_1}, -\sqrt{4-4z_1}), \\ P_4 &= (\sqrt{z_2}, -\sqrt{4-4z_2}), & P_8 &= (-\sqrt{z_2}, -\sqrt{4-4z_2}). \end{aligned}$$

Dado que nuestra función f satisface

$$f(x, y) = -f(-x, y) = -f(x, -y) = f(-x, -y),$$

basta evaluar f solamente en P_1 y P_2 . Tomando en cuenta que

$$f_1 := f(\sqrt{z_1} = 0,98178379, \sqrt{4-4z_1} = 0,3800831) = 0,062756644,$$

$$f_2 := f(\sqrt{z_2} = 0,48015105, \sqrt{4-4z_2} = 1,7543716) = -1,247706,$$

concluimos que

$$\begin{aligned} \min_{(x,y) \in \mathcal{E}} f(x, y) &= \min_{(x,y) \in \partial\mathcal{E}} f(x, y) = f(\sqrt{z_2}, \sqrt{4-4z_2}) = -1,247706, \\ \max_{(x,y) \in \mathcal{E}} f(x, y) &= \max_{(x,y) \in \partial\mathcal{E}} f(x, y) = f(-\sqrt{z_2}, \sqrt{4-4z_2}) = 1,247706. \end{aligned}$$

2.7. Aplicaciones de $\mathbb{R}^n$ en $\mathbb{R}^m$

Ahora consideraremos aplicaciones $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$. Tales aplicaciones mapean un punto $x = (x_1, \dots, x_n) \in D(f)$ a un punto $y = (y_1, \dots, y_m) \in B(f) \subset \mathbb{R}^m$. Cada coordenada y_i , $i = 1, \dots, m$, depende de $(x_1, \dots, x_n)$, es decir f es definida por m funciones:

$$\begin{aligned} f_1 : \mathbb{R}^n &\rightarrow \mathbb{R}, & D(f_1) &= D(f); & y_1 &= f_1(x_1, \dots, x_n), \\ & & & \vdots & & \end{aligned}$$

$$f_m : \mathbb{R}^n \rightarrow \mathbb{R}, \quad D(f_m) = D(f); \quad y_m = f_m(x_1, \dots, x_n).$$

Escribimos $f = (f_1, \dots, f_m)$, donde f_i es la i -ésima función de coordenadas.

Teorema 2.11. *Sea $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$. La función f es continua en $x^0 \in D(f)$ si y sólo si cada función f_i , $i = 1, \dots, m$, es continua en x^0 .*

Demostración. La aplicación f es continua en x^0 si y sólo si para cada sucesión $\{x^k\}_{k \in \mathbb{N}}$ tal que $x^k \in D(f)$ y $x^k \rightarrow x^0$ cuando $k \rightarrow \infty$ se tiene que $f(x^k) \rightarrow f(x^0)$, es decir,

$$(f_1(x^k), \dots, f_m(x^k)) \rightarrow (f_1(x^0), \dots, f_m(x^0)) \quad \text{cuando } k \rightarrow \infty.$$

Pero según el Teorema 1.4, esto sucede si y sólo si $f_i(x^k) \rightarrow f_i(x^0)$ cuando $k \rightarrow \infty$ para cada $i \in \{1, \dots, m\}$ fijo. Esto a su vez es válido si y sólo si todas las funciones f_i son continuas en x^0 . ■

Ejemplo 2.11. *Un ejemplo importante son las aplicaciones lineales, donde $D(f) = \mathbb{R}^n$ y existen constantes $a_{\mu\nu} \in \mathbb{R}$ tales que*

$$\begin{aligned} y_1 &= a_{11}x_1 + \dots + a_{1n}x_n, \\ &\vdots \end{aligned}$$

$$y_m = a_{m1}x_1 + \cdots + a_{mn}x_n.$$

Teorema 2.12. *Cada aplicación lineal es continua.*

Demostración. Tarea. ■

También para funciones $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ definiremos ahora el concepto de la diferenciabilidad. En analogía al caso de una función $f : \mathbb{R}^n \rightarrow \mathbb{R}$ aquí la diferenciabilidad en un punto x^0 significa que en una vecindad de x^0 podemos aproximar $f(x) - f(x^0)$ hasta un error de primer orden por una aplicación lineal.

Definición 2.10. *Sea $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ y x^0 un punto interior de $D(f)$. La función $f = (f_1, \dots, f_m)$ se llama diferenciable en x^0 si existen una matriz*

$$C = (c_{\mu\nu})_{\substack{\mu=1,\dots,m \\ \nu=1,\dots,n}} \in \mathbb{R}^{m \times n}$$

y una aplicación $f^0 : \mathbb{R}^n \rightarrow \mathbb{R}^m$ definida en una vecindad $U(x^0)$ de x^0 tales que para $\mu = 1, \dots, m$ se tiene que

1. $f_\mu^0(x^0) = \lim_{x \rightarrow x^0} f_\mu^0(x) = 0,$
2. $f_\mu(x) = f_\mu(x^0) + \sum_{\nu=1}^n c_{\mu\nu}(x_\nu - x_\nu^0) + d(x, x^0)f_\mu^0(x).$

Según esta definición, una función $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ es diferenciable en un punto si cada función coordenada f_i es diferenciable. Observamos que

$$c_{\mu\nu} = \frac{\partial f_\mu}{\partial x_\nu}(x^0), \quad \mu = 1, \dots, m, \quad \nu = 1, \dots, n. \quad (2.13)$$

Definición 2.11. *Sea $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$. Si existen las derivadas parciales $\partial f_\mu / \partial x_\nu(x^0)$ para $\mu = 1, \dots, m$ y $\nu = 1, \dots, n$ en un punto interior $x^0 \in D(f)$, entonces la matriz*

$$\frac{df}{dx}(x^0) = J_f(x^0) = \frac{\partial(f_1, \dots, f_m)}{\partial(x_1, \dots, x_n)}(x^0) = \left(\frac{\partial f_\mu}{\partial x_\nu}(x^0) \right)$$

se llama matriz funcional de Jacobi o matriz Jacobiana de f en x^0 .

Teorema 2.13. *Sea $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$. Si en un punto interior x^0 de $D(f)$ existe la matriz funcional de Jacobi $J_f(x^0)$, entonces existen un $\delta > 0$ y una constante M tal que para todo h con $|h| < \delta$ y $\nu = 1, \dots, n$ se tiene que*

$$d(f(x^0 + h\vec{e}_\nu), f(x^0)) \leq M|h|.$$

Demostración. Tarea. ■

Ejemplo 2.12. *Sea la función $f : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ definida por $f = (f_1, f_2, f_3)$ con*

$$f_1(x, y) = 2x + y, \quad f_2(x, y) = 3x^2 + y^2, \quad f_3(x, y) = xy.$$

Aquí obtenemos las derivadas parciales

$$\frac{\partial f_1}{\partial x} = 2, \quad \frac{\partial f_1}{\partial y} = 1; \quad \frac{\partial f_2}{\partial x} = 6x, \quad \frac{\partial f_2}{\partial y} = 2y; \quad \frac{\partial f_3}{\partial x} = y, \quad \frac{\partial f_3}{\partial y} = x.$$

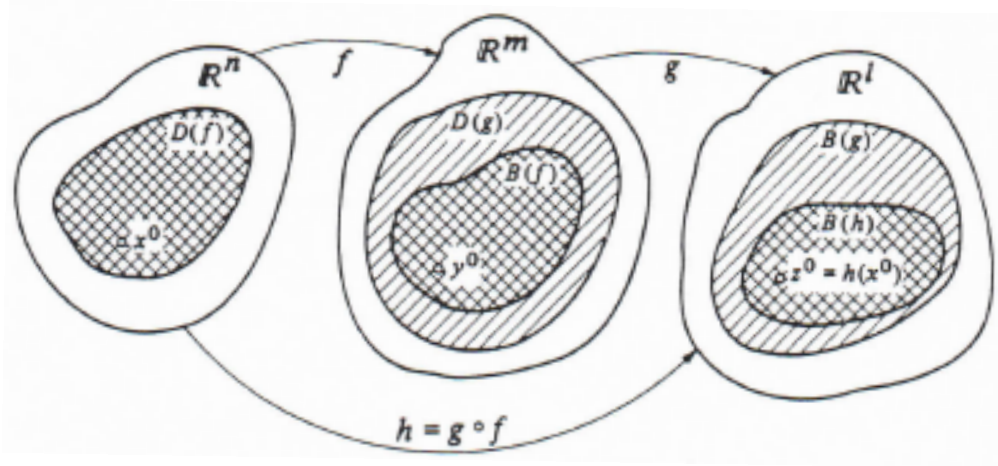


FIGURA 2.10. Ilustración del Teorema 2.14.

Para el punto $(1, -2)$ obtenemos

$$\frac{\partial(f_1, f_2, f_3)}{\partial(x, y)}(1, -2) = J_f((1, -2)) = \begin{bmatrix} 2 & 1 \\ 6 & -4 \\ -2 & 1 \end{bmatrix}.$$

Ejemplo 2.13. Sea la aplicación $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ definida por

$$f_1(x, y) = x \cos y, \quad f_2(x, y) = x \sin y.$$

Entonces

$$\frac{\partial(f_1, f_2)}{\partial(x, y)}(x, y) = J_f(x, y) = \begin{bmatrix} \cos y & -x \sin y \\ \sin y & x \cos y \end{bmatrix}.$$

2.8. La regla de la cadena

Teorema 2.14 (Regla de la cadena). Se consideran las dos funciones

$$f : \mathbb{R}^n \supset D(f) \ni x \mapsto f(x) \in B(f) \subset \mathbb{R}^m,$$

$$g : \mathbb{R}^m \supset D(g) \ni y \mapsto g(y) \in B(g) \subset \mathbb{R}^l,$$

donde sea $B(f) \subset D(g)$ y $h = g \circ f$. Además, sean x^0 un punto interior de $D(f)$ e $y^0 = f(x^0)$ un punto interior de $D(g)$. Sea la función g diferenciable en y^0 . En este caso se tiene lo siguiente.

1. Si $df/dx(x^0)$ existe, también existe $dh/dx(x^0)$, y

$$\frac{dh}{dx}(x^0) = \frac{dg}{dy}(y^0) \cdot \frac{df}{dx}(x^0).$$

2. Si f es diferenciable en x^0 , también h es diferenciable en x^0 .

Demostración. Sean $f = (f_1, \dots, f_m)$, $g = (g_1, \dots, g_l)$ y $h = (h_1, \dots, h_l)$. Puesto que g es diferenciable en y^0 , se tiene lo siguiente en una vecindad $U(y^0)$ de y^0 para $\lambda = 1, \dots, l$:

$$\begin{aligned} g_\lambda(y) - g_\lambda(y^0) &= \nabla g_\lambda(y^0) \cdot (\vec{y} - \vec{y}^0) + d(y, y^0)g_\lambda(y) \\ &= \sum_{\mu=1}^m \frac{\partial g_\lambda}{\partial y_\mu}(y^0)(y_\mu - y_\mu^0) + d(y, y^0)g_\lambda^0(y), \end{aligned} \quad (2.14)$$

donde $g_\lambda^0(y)$ es continua en y^0 y satisface $g_\lambda^0(y^0) = 0$.

1. Supongamos que $df/dx(x^0)$ existe. Para $\nu = 1, \dots, n$ definimos $x = x^0 + t\vec{e}_\nu$ para un $t \in \mathbb{R}$. Entonces existe un $t_0 > 0$ tal que $f(x^0 + t\vec{e}_\nu) \in U(y^0)$ para todo t con $|t| < t_0$ y todo $\nu = 1, \dots, n$. Ahora si insertamos lo siguiente en (2.14):

$$y = f(x^0 + t\vec{e}_\nu), \quad y_\mu = f_\mu(x^0 + t\vec{e}_\nu); \quad y^0 = f(x^0), \quad y_\mu^0 = f_\mu(x^0),$$

y tomamos en cuenta que $h_\lambda(x) = g_\lambda(f(x))$ para todo t con $0 < |t| < t_0$, obtenemos

$$\begin{aligned} \frac{h_\lambda(x^0 + t\vec{e}_\nu) - h_\lambda(x^0)}{t} &= \frac{g_\lambda(f(x^0 + t\vec{e}_\nu)) - g_\lambda(f(x^0))}{t} \\ &= \sum_{\mu=1}^m \frac{\partial g_\lambda}{\partial y_\mu}(y^0) \frac{f_\mu(x^0 + t\vec{e}_\nu) - f_\mu(x^0)}{t} \\ &\quad + \frac{d(f(x^0 + t\vec{e}_\nu), f(x^0))g_\lambda^0(f(x^0 + t\vec{e}_\nu))}{t}. \end{aligned}$$

Para $t \rightarrow 0$ tenemos las convergencias

$$\begin{aligned} \frac{f_\mu(x^0 + t\vec{e}_\nu) - f_\mu(x^0)}{t} &\rightarrow \frac{\partial f_\mu}{\partial x_\nu}(x^0), \\ g_\lambda(f(x^0 + t\vec{e}_\nu)) &\rightarrow g_\lambda^0(f(x^0)) = g_\lambda^0(y^0) = 0, \end{aligned}$$

por lo tanto en virtud del Teorema 2.13,

$$\frac{\partial h_\lambda}{\partial x_\nu} = \lim_{t \rightarrow 0} \frac{h_\lambda(x^0 + t\vec{e}_\nu) - h_\lambda(x^0)}{t} = \sum_{\mu=1}^m \frac{\partial g_\lambda}{\partial y_\mu}(y^0) \cdot \frac{\partial f_\mu}{\partial x_\nu}(x^0).$$

Tomando en cuenta la definición del producto entre dos matrices, obtenemos el primer enunciado del teorema.

2. Si f es diferenciable en x^0 , entonces se tiene lo siguiente en una vecindad de x^0 para $\mu = 1, \dots, m$:

$$\begin{aligned} f_\mu(x) - f_\mu(x^0) &= \nabla f_\mu(x^0) \cdot (\vec{x} - \vec{x}^0) + d(x, x^0)f_\mu^0(x) \\ &= \sum_{\nu=1}^n \frac{\partial f_\mu}{\partial x_\nu}(x^0)(x_\nu - x_\nu^0) + d(x, x^0)f_\mu^0(x), \end{aligned}$$

donde las funciones $f_\mu^0(x)$ son continuas en x^0 y satisfacen $f_\mu^0(x^0) = 0$. Insertando esto en (2.14) entrega que

$$g_\lambda(f(x)) - g_\lambda(f(x^0)) = \sum_{\mu=1}^m \frac{\partial g_\lambda}{\partial y_\mu}(y^0) \left(\sum_{\nu=1}^n \frac{\partial f_\mu}{\partial x_\nu}(x^0)(x_\nu - x_\nu^0) + d(x, x^0)f_\mu^0(x) \right)$$

$$+ d(f(x), f(x^0))g_\lambda^0(f(x)).$$

En virtud de lo anterior,

$$h_\lambda(x) = h_\lambda(x^0) + \sum_{\nu=1}^n \left(\sum_{\mu=1}^m \frac{\partial g_\lambda}{\partial y_\mu}(y^0) \frac{\partial f_\mu}{\partial x_\nu}(x^0) \right) (x_\nu - x_\nu^0) + d(x, x^0)h_\lambda^0(x),$$

donde definimos

$$h_\lambda^0(x) = \begin{cases} \sum_{\mu=1}^m \frac{\partial g_\lambda}{\partial y_\mu}(y^0) f_\mu^0(x) + \frac{d(f(x), f(x^0))}{d(x, x^0)} g_\lambda^0(f(x)) & \text{si } x \neq x^0, \\ 0 & \text{si } x = x^0. \end{cases}$$

Según el Teorema 2.6 existen una constante $M > 0$ y un $\delta > 0$ tales que para todo x con $0 < d(x, x^0) < \delta$ se tiene que

$$\frac{d(f(x), f(x^0))}{d(x, x^0)} = \sqrt{\sum_{\mu=1}^m \left\{ \frac{f_\mu(x) - f_\mu(x^0)}{d(x, x^0)} \right\}^2} \leq M.$$

Puesto que

$$\lim_{x \rightarrow x^0} f_\mu^0(x) = 0, \quad \lim_{x \rightarrow x^0} g_\lambda^0(f(x)) = g_\lambda^0(y^0) = 0$$

se tiene que

$$\lim_{x \rightarrow x^0} h_\lambda^0(x) = 0 = h_\lambda^0(x^0).$$

Entonces cada una de las funciones h_λ es diferenciable en x^0 , lo que concluye la demostración del teorema. ■

Comentamos que la ecuación que aparece en el Teorema 2.14,

$$\frac{dh}{dx}(x^0) = \frac{dg}{dy}(y^0) \cdot \frac{df}{dx}(x^0)$$

es una ecuación matricial de la forma

$$\begin{bmatrix} \frac{\partial h_1}{\partial x_1} & \frac{\partial h_1}{\partial x_2} & \cdots & \frac{\partial h_1}{\partial x_n} \\ \frac{\partial h_2}{\partial x_1} & \frac{\partial h_2}{\partial x_2} & \cdots & \frac{\partial h_2}{\partial x_n} \\ \vdots & & \ddots & \vdots \\ \frac{\partial h_l}{\partial x_1} & \frac{\partial h_l}{\partial x_2} & \cdots & \frac{\partial h_l}{\partial x_n} \end{bmatrix} = \begin{bmatrix} \frac{\partial g_1}{\partial y_1} & \frac{\partial g_1}{\partial y_2} & \cdots & \frac{\partial g_1}{\partial y_m} \\ \frac{\partial g_2}{\partial y_1} & \frac{\partial g_2}{\partial y_2} & \cdots & \frac{\partial g_2}{\partial y_m} \\ \vdots & & \ddots & \vdots \\ \frac{\partial g_l}{\partial y_1} & \frac{\partial g_l}{\partial y_2} & \cdots & \frac{\partial g_l}{\partial y_m} \end{bmatrix} \cdot \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_n} \\ \vdots & & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \frac{\partial f_m}{\partial x_2} & \cdots & \frac{\partial f_m}{\partial x_n} \end{bmatrix},$$

donde la primera y la tercera matriz deben ser evaluadas en x^0 y la segunda en $y^0 = f(x^0)$.

Por otro lado, en las aplicaciones frecuentemente se presenta la situación $f: \mathbb{R} \rightarrow \mathbb{R}^m$ y $g: \mathbb{R}^m \rightarrow \mathbb{R}$. En este caso, la aplicación $h = g \circ f: \mathbb{R} \rightarrow \mathbb{R}$ esta definida por

$$h(x) = g(y_1, \dots, y_m),$$

donde

$$y_1 = f_1(x), \dots, y_m = f_m(x).$$

Aquí la regla de la cadena asume la forma

$$\frac{dh}{dx}(x^0) = \begin{bmatrix} \frac{\partial g}{\partial y_1}(y^0) & \dots & \frac{\partial g}{\partial y_m}(y^0) \end{bmatrix} \cdot \begin{bmatrix} \frac{df_1}{dx}(x^0) \\ \vdots \\ \frac{df_m}{dx}(x^0) \end{bmatrix} = \sum_{\mu=1}^m \frac{\partial g}{\partial y_\mu}(y^0) \cdot \frac{df_\mu}{dx}(x^0).$$

Ejemplo 2.14. Sea

$$h(x) = g(y_1, y_2) = e^{y_1 y_2}, \quad y_1 = f_1(x) = x \cos x, \quad y_2 = f_2(x) = x \sin x.$$

Queremos calcular $h'(x^0)$ para $x^0 = \pi$. Poniendo

$$y^0 = (y_1^0, y_2^0) = (f_1(x^0), f_2(x^0)) = (-\pi, 0)$$

obtenemos

$$\frac{\partial g}{\partial y_1} = y_2 e^{y_1 y_2}, \quad \frac{\partial g}{\partial y_1}(y_1^0, y_2^0) = 0; \quad \frac{\partial g}{\partial y_2} = y_1 e^{y_1 y_2}, \quad \frac{\partial g}{\partial y_2}(y_1^0, y_2^0) = -\pi;$$

además se tiene que

$$\frac{df_1}{dx} = \cos x - x \sin x, \quad \frac{df_1}{dx}(x^0) = -1; \quad \frac{df_2}{dx} = \sin x + x \cos x; \quad \frac{df_2}{dx} = -\pi.$$

Estos valores nos permiten calcular

$$h'(\pi) = \frac{\partial g}{\partial y_1}(y_1^0, y_2^0) \frac{df_1}{dx}(x^0) + \frac{\partial g}{\partial y_2}(y_1^0, y_2^0) \frac{df_2}{dx}(x^0) = \pi^2.$$

Ejemplo 2.15. La regla de la cadena se usa mucho para derivar funciones que provienen de otras funciones a través de una sustitución de variables. Por ejemplo, sea dada la función $f(x, y)$. Introduciendo las llamadas coordenadas polares planas

$$x = r \cos \varphi, \quad y = r \sin \varphi$$

podemos convertir $f(x, y)$ en una función $F(r, \varphi)$:

$$F(r, \varphi) = f(r \cos \varphi, r \sin \varphi).$$

La regla de la cadena nos entrega la siguiente relación entre $\partial F / \partial r$, $\partial F / \partial \varphi$, $\partial f / \partial x$ y $\partial f / \partial y$:

$$\begin{aligned} \frac{\partial F}{\partial r} &= \frac{\partial f}{\partial x} \frac{\partial}{\partial r}(r \cos \varphi) + \frac{\partial f}{\partial y} \frac{\partial}{\partial r}(r \sin \varphi) = \frac{\partial f}{\partial x} \cos \varphi + \frac{\partial f}{\partial y} \sin \varphi, \\ \frac{\partial F}{\partial \varphi} &= \frac{\partial f}{\partial x} \frac{\partial}{\partial \varphi}(r \cos \varphi) + \frac{\partial f}{\partial y} \frac{\partial}{\partial \varphi}(r \sin \varphi) = -\frac{\partial f}{\partial x} r \sin \varphi + \frac{\partial f}{\partial y} r \cos \varphi. \end{aligned}$$

2.9. El Teorema del Valor Intermedio

Definición 2.12. Sean $x^1, x^2 \in \mathbb{R}^n$, $x^1 \neq x^2$. Entonces al siguiente conjunto se dice segmento lineal que une los puntos x^1 y x^2 :

$$\overline{x^1, x^2} = \{x \in \mathbb{R}^n \mid x = x^1 + \lambda(\vec{x}^2 - \vec{x}^1), \lambda \in [0, 1]\}.$$

Teorema 2.15 (Teorema del Valor Intermedio). Sea $f : \mathbb{R}^n \rightarrow \mathbb{R}$. Sea $X \subset D(f)$ un conjunto abierto, y sea f diferenciable sobre X . Sean $x^1, x^2 \in X$, $x^1 \neq x^2$ y sea el segmento lineal $\overline{x^1, x^2}$ contenido en X . Entonces existe por lo menos un $\xi \in \overline{x^1, x^2}$ tal que $\xi \neq x^1$, $\xi \neq x^2$, y

$$f(x^2) - f(x^1) = \nabla f(\xi) \cdot (\vec{x}^2 - \vec{x}^1).$$

Demostración. Consideremos para $t \in [0, 1]$ la función $F : \mathbb{R} \rightarrow \mathbb{R}$ definida por

$$F(t) := f(x^1 + t(\vec{x}^2 - \vec{x}^1)).$$

Esta función es continua sobre $[0, 1]$ y diferenciable sobre $(0, 1)$, así que según el Teorema del Valor Intermedio de funciones de una variable existe un $\vartheta \in (0, 1)$ tal que

$$F(1) - F(0) = F'(\vartheta).$$

Según el Teorema 2.14,

$$F'(t) = \left(\nabla f(x^1 + t(\vec{x}^2 - \vec{x}^1)) \right) \cdot (\vec{x}^2 - \vec{x}^1),$$

por lo tanto

$$F(1) - F(0) = \left(\nabla f(x^1 + \vartheta(\vec{x}^2 - \vec{x}^1)) \right) \cdot (\vec{x}^2 - \vec{x}^1).$$

Para $\xi = x^1 + \vartheta(\vec{x}^2 - \vec{x}^1)$ se tiene que

$$f(x^2) - f(x^1) = F(1) - F(0) = \nabla f(\xi) \cdot (\vec{x}^2 - \vec{x}^1).$$

■

Un caso particular de los conjuntos X que satisfacen las hipótesis del Teorema 2.15 son los conjuntos *convexos*.

Definición 2.13. Un conjunto $X \subset \mathbb{R}^n$ se llama convexo si con cada par de puntos x^1 y x^2 también el segmento lineal $\overline{x^1, x^2}$ está contenido en X .

Comentamos que para $n = 1$, el enunciado del Teorema 2.15 es

$$f(x^2) - f(x^1) = f'(\xi)(x^2 - x^1),$$

lo que es el Teorema del Valor Intermedio para funciones de una variable. Para funciones $f(x, y)$ de dos variables, el Teorema 2.15 entrega la siguiente fórmula, donde definimos $x^1 = (x_1, y_1)$ y $x^2 = (x_2, y_2)$:

$$f(x_2, y_2) - f(x_1, y_1) = (x_2 - x_1)f_x(\xi, \eta) + (y_2 - y_1)f_y(\xi, \eta),$$

donde ξ está localizado entre x_1 y x_2 y η entre y_1 e y_2 .

Finalmente, comentamos que no existe una versión del Teorema 2.15 válida para funciones $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$. Para ver eso, basta discutir el ejemplo $f : \mathbb{R} \rightarrow \mathbb{R}^2$ con $f(x) = (x^2, x^3)$ sobre $[0, 1]$.

Para funciones de una variable ya sabemos que del Teorema del Valor Intermedio sigue que una función f continua sobre $[a, b]$ y diferenciable sobre (a, b) es constante si y sólo si $f'(x) = 0$ sobre (a, b) . Podemos definir un teorema análogo para funciones de n variables basándonos en el concepto de una *región* en $\mathbb{R}^n$.

Definición 2.14. Un conjunto $G \subset \mathbb{R}^n$ se llama

1. *poligonalmente conexo* si para cada par $x, \tilde{x} \in G$ existe un número finito de puntos $x = x^1, x^2, \dots, x^m = \tilde{x} \in G$ tales que el trazado poligonal

$$P = \bigcup_{i=1}^{m-1} \overline{x^i, x^{i+1}}$$

pertenece enteramente a G ;

2. *una región* si G es abierto y poligonalmente conexo.

Teorema 2.16. Sea $f : \mathbb{R}^n \rightarrow \mathbb{R}$; además, sea $G \subset \mathbb{R}^n$ una región y sea f diferenciable sobre G . Entonces f es constante sobre G si y sólo si

$$\nabla f(x) = \vec{0} \quad \text{para todo } x \in G.$$

Demostración.

1. Si f es constante sobre G , entonces $\nabla f(x) = \vec{0}$ para todo $x \in G$.
2. Sea $\nabla f(x) = \vec{0}$ para todo $x \in G$. Si $x, \tilde{x} \in G$, entonces existen puntos $x^1 = x, x^2, \dots, x^m = \tilde{x}$ tales que los segmentos lineales $\overline{x^i, x^{i+1}}$ para $i = 1, \dots, m-1$ están completamente contenidos en G . Según el Teorema 2.15 existe un $\xi^i \in \overline{x^i, x^{i+1}}$ para el cual

$$f(x^{i+1}) - f(x^i) = \nabla f(\xi^i) \cdot (\vec{x}^{i+1} - \vec{x}^i) = 0,$$

por lo tanto

$$f(x^i) = f(x^{i+1}) \quad \text{para } i = 1, \dots, m-1$$

y luego $f(x) = f(\tilde{x})$. Dado x fue elegido arbitrario, concluimos que $f(x) = f(\tilde{x})$ para todo $x \in G$, es decir f es constante sobre G . ■

Ejemplo 2.16. Consideremos la función $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ definida por

$$f(x, y) = \arctan \frac{x}{y} + \arctan \frac{y}{x}$$

sobre las regiones

$$G_1 = \{(x, y) \mid x > 0, y > 0\},$$

$$G_2 = \{(x, y) \mid x < 0, y > 0\},$$

$$G_3 = \{(x, y) \mid x < 0, y < 0\},$$

$$G_3 = \{(x, y) \mid x > 0, y < 0\}.$$

La función f es diferenciable sobre $G = G_1 \cup G_2 \cup G_3 \cup G_4$, y se tiene que

$$\frac{\partial f}{\partial x} = \frac{1}{1 + \left(\frac{y}{x}\right)^2} \cdot \left(-\frac{y}{x^2}\right) + \frac{1}{1 + \left(\frac{x}{y}\right)^2} \cdot \frac{1}{y} = 0,$$

$$\frac{\partial f}{\partial y} = \frac{1}{1 + \left(\frac{y}{x}\right)^2} \cdot \frac{1}{x} + \frac{1}{1 + \left(\frac{x}{y}\right)^2} \cdot \left(-\frac{x}{y^2}\right) = 0,$$

por lo tanto $\nabla f(x, y) = \vec{0}$ sobre $G_1 \cup G_2 \cup G_3 \cup G_4$. Puesto que cada uno de los conjuntos G_1 , G_2 , G_3 y G_4 es una región, el Teorema 2.16 implica que $f(x, y)$ debe ser constante sobre cada una de estas regiones. Efectivamente,

$$\begin{aligned} f(x, y) &\equiv f(1, 1) = 2 \arctan 1 = \frac{\pi}{2} && \text{para todo } (x, y) \in G_1, \\ f(x, y) &\equiv f(-1, 1) = 2 \arctan(-1) = -\frac{\pi}{2} && \text{para todo } (x, y) \in G_2, \\ f(x, y) &\equiv f(-1, -1) = 2 \arctan 1 = \frac{\pi}{2} && \text{para todo } (x, y) \in G_3, \\ f(x, y) &\equiv f(1, -1) = 2 \arctan(-1) = -\frac{\pi}{2} && \text{para todo } (x, y) \in G_4. \end{aligned}$$

Note que no podemos aplicar el Teorema 2.16 para el conjunto G , puesto que G es un conjunto abierto, pero no es poligonalmente conexo y por lo tanto no es una región. Tenemos $\nabla f(x, y) = \vec{0}$ para todo $(x, y) \in G$, pero $f(x, y)$ no es constante sobre G . Solamente sobre cada una de las cuatro sub-regiones $f(x, y)$ es constante.

Finalmente comunicamos los siguientes teoremas (sin demostraciones) que establecen las relaciones entre un conjunto conexo en el sentido de la Definición 1.12 y un conjunto poligonalmente conexo.

Teorema 2.17. Sea $G \subset \mathbb{R}^n$ un conjunto poligonalmente conexo. Entonces G es conexo.

Teorema 2.18. Un conjunto $G \subset \mathbb{R}^n$ es una región si y sólo si es abierto y conexo.

Capítulo 3

Funciones de $\mathbb{R}^n$ en $\mathbb{R}^m$ II: aplicaciones

3.1. El Teorema de Taylor

Recordemos primeramente el Teorema de Taylor para funciones de una variable (Cálculo I). Sea $f : \mathbb{R} \rightarrow \mathbb{R}$. Sea I un intervalo arbitrario, sean $x_0 \in I$ y $m \in \mathbb{N}_0$, y sea la función f por lo menos $m + 1$ veces continuamente diferenciable sobre I . Entonces para todo $x \in I$ es válida la *fórmula de Taylor*

$$f(x) = T_m(x) + R_m(x, x_0)$$

con el polinomio de Taylor

$$T_m(x) = \sum_{k=0}^m \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k$$

y el término residual en su forma de Lagrange:

$$R_m(x, x_0) = \frac{f^{(m+1)}(x_0 + \vartheta(x - x_0))}{(m+1)!} (x - x_0)^{m+1}$$

para un $\vartheta \in (0, 1)$.

Trataremos ahora una generalización del Teorema de Taylor para funciones $f : \mathbb{R}^n \rightarrow \mathbb{R}$. Para tal efecto definimos ahora el concepto de un polinomio de n variables.

Definición 3.1. Un polinomio de n variables es una función $p : \mathbb{R}^n \rightarrow \mathbb{R}$ con $D(p) = \mathbb{R}^n$ y

$$p(x_1, \dots, x_n) = \sum_{\nu_1, \nu_2, \dots, \nu_n=0}^N a_{\nu_1 \nu_2 \dots \nu_n} x_1^{\nu_1} x_2^{\nu_2} \dots x_n^{\nu_n},$$

donde $\nu_1, \dots, \nu_n, N \in \mathbb{N}_0$ y $a_{\nu_1 \nu_2 \dots \nu_n} \in \mathbb{R}$. El mayor de los números $\nu_1 + \nu_2 + \dots + \nu_n$ con $a_{\nu_1 \nu_2 \dots \nu_n} \neq 0$ se llama el grado de p .

Definición 3.2. Sea $\vec{h} = \{h_1, \dots, h_n\}$ un vector de $\mathbb{V}^n$ y $k \in \mathbb{N}$. Sea $f : \mathbb{R}^n \rightarrow \mathbb{R}$ y sea $f \in C^k(X)$ para un conjunto $X \subset D(f)$. Entonces definimos el operador diferencial $(\vec{h} \cdot \nabla)^k$ mediante

$$(\vec{h} \cdot \nabla)^k(x) = \sum_{\nu_1, \nu_2, \dots, \nu_k=1}^n h_{\nu_1} h_{\nu_2} \dots h_{\nu_k} \frac{\partial^k f}{\partial x_{\nu_1} \partial x_{\nu_2} \dots \partial x_{\nu_k}}(x)$$

para $k > 0$ y formalmente

$$(\vec{h} \cdot \nabla)^0(x) = f(x).$$

Ejemplo 3.1. Consideremos el caso $n = 2$, $k = 2$. En este caso el operador diferencial $(\vec{h} \cdot \nabla)^2$ para una función $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ sobre un conjunto $X \subset D(f)$ con $f \in C^2(X)$ está (según la Definición 3.2) dado por

$$\begin{aligned} (\vec{h} \cdot \nabla f)^2(x) &= \left(h_1 \frac{\partial f}{\partial x_1} + h_2 \frac{\partial f}{\partial x_2} \right)^2 (x_1, x_2) \\ &= h_1 h_1 \frac{\partial^2 f}{\partial x_1 \partial x_1} (x_1, x_2) + h_1 h_2 \frac{\partial^2 f}{\partial x_1 \partial x_2} (x_1, x_2) \\ &\quad + h_2 h_1 \frac{\partial^2 f}{\partial x_2 \partial x_1} (x_1, x_2) + h_2 h_2 \frac{\partial^2 f}{\partial x_2 \partial x_2} (x_1, x_2) \\ &= h_1^2 \frac{\partial^2 f}{\partial x_1^2} (x_1, x_2) + 2h_1 h_2 \frac{\partial^2 f}{\partial x_1 \partial x_2} (x_1, x_2) + h_2^2 \frac{\partial^2 f}{\partial x_2^2} (x_1, x_2), \end{aligned}$$

donde utilizamos la siguiente relación (la cual es válida según el Teorema 2.9, puesto que $f \in C^2(X)$):

$$\frac{\partial^2 f}{\partial x_1 \partial x_2} = \frac{\partial^2 f}{\partial x_2 \partial x_1}.$$

Teorema 3.1 (Teorema de Taylor). Sea $f : \mathbb{R}^n \rightarrow \mathbb{R}$ y $X \subset D(f)$, y para un $m \in \mathbb{N}_0$ sea $f \in C^{m+1}(X)$. Además sean $x^0 \in X$ y $x \in X$, y sea el segmento lineal x^0, x contenido en X . Entonces los siguientes enunciados se tienen para $\vec{h} = \vec{x} - \vec{x}^0$:

1. La función f puede ser representada por la fórmula de Taylor

$$f(x) = \sum_{k=0}^m \frac{(\vec{h} \cdot \nabla f)^k(x^0)}{k!} + R_m(x, x^0).$$

2. El término residual $R_m(x, x^0)$ es de la siguiente forma, donde $\vartheta \in (0, 1)$:

$$R_m(x, x^0) = \frac{(\vec{h} \cdot \nabla f)^{m+1}(x^0 + \vartheta \vec{h})}{(m+1)!}.$$

Demostración. Note que el conjunto X es abierto en virtud de la Definición 2.9. Para $t \in [0, 1]$ consideramos la función $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ con

$$\varphi(t) = f(x^0 + t\vec{h}).$$

Esta función es $m+1$ veces continuamente diferenciable sobre $[0, 1]$, y según el Teorema de Taylor para funciones de una variable existe un $\vartheta \in (0, 1)$ tal que

$$\varphi(1) = \sum_{k=0}^m \frac{\varphi^{(k)}(0)}{k!} \cdot 1^k + \frac{\varphi^{(m+1)}(\vartheta)}{(m+1)!}. \quad (3.1)$$

Ahora la regla de la cadena entrega las siguientes derivadas de la función φ :

$$\begin{aligned} \varphi^{(0)}(t) &= f(x^0 + t\vec{h}) = (\vec{h} \cdot \nabla f)^0(x^0 + t\vec{h}), \\ \varphi'(t) &= \sum_{i=1}^n h_i \frac{\partial f(x^0 + t\vec{h})}{\partial x_i} = (\vec{h} \cdot \nabla f)(x^0 + t\vec{h}), \end{aligned}$$

$$\varphi''(t) = \sum_{i=1}^n h_i \left(\sum_{j=1}^n h_j \frac{\partial^2 f(x^0 + t\vec{h})}{\partial x_i \partial x_j} \right) = (\vec{h} \cdot \nabla f)^2(x^0 + t\vec{h}), \quad \text{etc.}$$

En general obtenemos

$$\varphi^{(k)}(t) = (\vec{h} \cdot \nabla f)^k(x^0 + t\vec{h}), \quad k = 0, 1, \dots, m+1,$$

y por lo tanto

$$\begin{aligned} \varphi^{(k)}(0) &= (\vec{h} \cdot \nabla f)^k(x^0), \quad k = 0, 1, \dots, m, \\ \varphi^{(m+1)} &= (\vec{h} \cdot \nabla f)^{m+1}(x^0 + \vartheta\vec{h}). \end{aligned}$$

Insertando esto en (3.1) obtenemos el enunciado del teorema. ■

En el caso $m = 0$, es decir, para $f \in C^1(X)$, obtenemos para un $\vartheta \in (0, 1)$ apropiado

$$f(x) = f(x^0) + \vec{h} \cdot \nabla f(x^0 + \vartheta\vec{h}).$$

Esto es el enunciado del Teorema del Valor Intermedio. Entonces, también aquí el Teorema de Taylor es una generalización del Teorema del Valor Intermedio.

Para una función $f(x, y)$ definida en $\mathbb{R}^2$ y $m = 1$, la fórmula de Taylor escrita detalladamente es la siguiente:

$$\begin{aligned} &f(x_0 + h, y_0 + k) \\ &= \frac{1}{0!} f(x_0, y_0) + \frac{1}{1!} \left(h \frac{\partial f}{\partial x} + k \frac{\partial f}{\partial y} \right)^1 (x_0, y_0) + \frac{1}{2!} \left(h \frac{\partial f}{\partial x} + k \frac{\partial f}{\partial y} \right)^2 (x_0 + \vartheta h, y_0 + \vartheta k) \\ &= f(x_0, y_0) + h f_x(x_0, y_0) + k f_y(x_0, y_0) \\ &\quad + \frac{1}{2} (h^2 f_{xx}(x_0 + \vartheta h, y_0 + \vartheta k) + 2hk f_{xy}(x_0 + \vartheta h, y_0 + \vartheta k) + k^2 f_{yy}(x_0 + \vartheta h, y_0 + \vartheta k)). \end{aligned}$$

Ejemplo 3.2 (Problema 5, Certamen 1, Curso 2009/I). Sea la función $f : \mathbb{R}^3 \rightarrow \mathbb{R}$ dada por

$$f(x, y, z) = e^x \sin y \cos z.$$

- Determinar el polinomio de Taylor $T_2(x, y, z)$ de f correspondiente al punto de desarrollo $(x_0, y_0, z_0) = (0, 0, 0)$.
- Determinar el termino residual $R_2(x, y, z, x_0, y_0, z_0)$ correspondiente.
- Sea

$$B := \{(x, y, z) \in \mathbb{R}^3 \mid -1 \leq x \leq 1, -\pi/4 \leq y \leq \pi/4, -\pi/4 \leq z \leq \pi/4\}.$$

Determinar una constante $C > 0$ tal que

$$\max_{(x, y, z) \in B} |R_2(x, y, z, x_0, y_0, z_0)| \leq C.$$

Solución sugerida.

a) Aquí calculamos las derivadas en $(0, 0, 0)$

$$\begin{aligned} f &= 0, & f_{xy} &= e^x \cos y \cos z \Big|_{(0,0,0)} = 1, \\ f_x &= e^x \sin y \cos z \Big|_{(0,0,0)} = 0, & f_{xz} &= -e^x \sin y \sin z \Big|_{(0,0,0)} = 0, \\ f_y &= e^x \cos y \cos z \Big|_{(0,0,0)} = 1, & f_{yz} &= -e^x \cos y \sin z \Big|_{(0,0,0)} = 0, \\ f_z &= -e^x \sin y \sin z \Big|_{(0,0,0)} = 0, & f_{yy} &= -e^x \sin y \cos z \Big|_{(0,0,0)} = 0, \\ f_{xx} &= e^x \sin y \cos z \Big|_{(0,0,0)} = 0, & f_{zz} &= -e^x \sin y \cos z \Big|_{(0,0,0)} = 0, \end{aligned}$$

por lo tanto

$$\begin{aligned} T_2(x, y, z) &= f(0, 0, 0) + x \frac{\partial f}{\partial x}(0, 0, 0) + y \frac{\partial f}{\partial y}(0, 0, 0) + z \frac{\partial f}{\partial z}(0, 0, 0) \\ &\quad + \frac{1}{2} \left[x^2 \frac{\partial^2 f}{\partial x^2}(0, 0, 0) + 2xy \frac{\partial^2 f}{\partial x \partial y}(0, 0, 0) + 2xz \frac{\partial^2 f}{\partial x \partial z}(0, 0, 0) \right. \\ &\quad \left. + 2yz \frac{\partial^2 f}{\partial y \partial z}(0, 0, 0) + y^2 \frac{\partial^2 f}{\partial y^2}(0, 0, 0) + z^2 \frac{\partial^2 f}{\partial z^2}(0, 0, 0) \right] \\ &= y + \frac{1}{2} \cdot 2xy = y + xy. \end{aligned}$$

b) En una notación simplificada obtenemos aquí las terceras derivadas

$$\begin{aligned} f_{xxx} &= e^x \sin y \cos z, & f_{xzz} &= -e^x \sin y \cos z, \\ f_{xxy} &= e^x \cos y \cos z, & f_{yyy} &= -e^x \cos y \cos z, \\ f_{xxz} &= -e^x \sin y \sin z, & f_{yyz} &= e^x \sin y \sin z, \\ f_{xyy} &= -e^x \sin y \cos z, & f_{yzz} &= -e^x \cos y \cos z, \\ f_{xyz} &= -e^x \cos y \sin z, & f_{zzz} &= e^x \sin y \sin z. \end{aligned}$$

El término residual es, en este caso,

$$\begin{aligned} R_2(x, y, z, x_0, y_0, z_0) &= \frac{1}{6} (x^3 f_{xxx} + y^3 f_{yyy} + z^3 f_{zzz} + 3x^2 y f_{xxy} + 3xy^2 f_{xyy} \\ &\quad + 3x^2 z f_{xxz} + 3xz^2 f_{xzz} + 3y^2 z f_{yyz} + 3yz^2 f_{yzz} + 6xyz f_{xyz}), \end{aligned}$$

donde las terceras derivadas son evaluadas en algún punto (ξ, η, ζ) con $0 \leq \xi \leq x$, $0 \leq \eta \leq y$ y $0 \leq \zeta \leq z$, es decir,

$$\begin{aligned} R_2(x, y, z, x_0, y_0, z_0) &= \frac{e^\xi}{6} (x^3 \sin \eta \cos \zeta - y^3 \cos \eta \cos \zeta + z^3 \sin \eta \sin \zeta \\ &\quad + 3x^2 y \cos \eta \cos \zeta - 3xy^2 \sin \eta \cos \zeta - 3x^2 z \sin \eta \sin \zeta \\ &\quad - 3xz^2 \sin \eta \cos \zeta + 3y^2 z \sin \eta \sin \zeta - 3yz^2 \cos \eta \cos \zeta \\ &\quad - 6xyz \cos \eta \sin \zeta). \end{aligned}$$

c) Tomando en cuenta que $|\sin \alpha| \leq 1/\sqrt{2}$ para $-\pi/4 \leq \alpha \leq \pi/4$, podemos acotar R_2 sobre B de la siguiente forma:

$$\begin{aligned} & |R_2(x, y, z, x_0, y_0, z_0)| \\ & \leq \frac{e}{6} \left(\frac{1}{\sqrt{2}} + \frac{\pi^3}{64} + \frac{\pi^3}{128} + \frac{3\pi}{4} + \frac{3\pi^2}{16\sqrt{2}} + \frac{3\pi}{8} + \frac{3\pi^2}{16\sqrt{2}} + \frac{3\pi^3}{128} + \frac{3\pi^3}{64} + \frac{6\pi^2}{16\sqrt{2}} \right) \\ & = \frac{e}{6} \left(\frac{1}{\sqrt{2}} + \frac{9\pi}{8} + \frac{(9\sqrt{2}+6)\pi^2}{32} + \frac{3\pi^3}{32} \right) \leq 5,855. \end{aligned}$$

3.2. Funciones implícitas

Queremos ahora estudiar la solución de sistemas de ecuaciones, empezando con elementos de la teoría de sistemas lineales. Sea $n > m$ y un sistema lineal dado por

$$\begin{aligned} a_{11}x_1 + \cdots + a_{1n}x_n - b_1 &= 0, \\ &\vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n - b_m &= 0. \end{aligned}$$

Suponemos que el rango de la matriz $(a_{\mu\nu})$ es m , es decir, después de un posible cambio de numeración de las variables,

$$\det(a_{\mu\nu})_{\mu,\nu=1,\dots,m} = \begin{vmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mm} \end{vmatrix} \neq 0. \quad (3.2)$$

En este caso podemos despejar $x_1, \dots, x_m$ del sistema lineal, es decir, existen coeficientes c_μ y $c_{\mu\nu}$ tales que

$$x_1 = c_1 + \sum_{\nu=m+1}^n c_{1\nu}x_\nu, \dots, x_m = c_m + \sum_{\nu=m+1}^n c_{m\nu}x_\nu.$$

Queremos generalizar ahora este problema y consideremos en $\mathbb{R}$ un sistema de ecuaciones

$$\begin{aligned} f_1(x) = f_1(x_1, \dots, x_n) &= 0, \\ &\vdots \\ f_m(x) = f_m(x_1, \dots, x_n) &= 0, \end{aligned}$$

donde suponemos que las funciones f_μ están todas definidas en una vecindad n -dimensional $U(x^0)$ de un punto x^0 , es decir,

$$U(x^0) \subset \bigcap_{\mu=1}^m D(f_\mu).$$

Ahora buscamos todos los puntos $x \in U(x^0)$ que son soluciones del sistema de ecuaciones. El siguiente ejemplo demuestra que un tal punto no necesariamente debe existir.

Ejemplo 3.3. *Consideremos el sistema*

$$\begin{aligned} f_1(x_1, x_2, x_3) &= x_1^2 + 2x_2^2 - 1 = 0, \\ f_2(x_1, x_2, x_3) &= x_1^2 + x_3^2 + 1 = 0. \end{aligned}$$

Aquí $n = 3$, $m = 2$. El sistema de ecuaciones no está satisfecho para ningún punto.

Para avanzar tenemos que exigir que por lo menos para un punto $x \in U(x^0)$ el sistema no lineal esté satisfecho. Además necesitaremos alguna otra condición que se convierte en la condición (3.2) para el caso especial de un sistema lineal.

Teorema 3.2 (Teorema Principal de Funciones Implícitas). *Sean dadas las funciones $f_\mu : \mathbb{R}^n \rightarrow \mathbb{R}$ para $\mu = 1, \dots, m$ con $m < n$. Sean $X \subset \mathbb{R}^n$ un conjunto abierto; $x^0 = (x_1^0, \dots, x_n^0) \in X$; $f_\mu \in C^1(X)$ para $\mu = 1, \dots, m$, y*

$$\begin{aligned} f_1(x_1^0, \dots, x_n^0) &= 0, \\ &\vdots \\ f_m(x_1^0, \dots, x_n^0) &= 0. \end{aligned}$$

Además, sea la matriz

$$F(x^0) = \left(\frac{\partial f_\mu}{\partial x_\nu}(x^0) \right)_{\mu, \nu=1, \dots, m} \quad (3.3)$$

no singular. En este caso, el sistema de ecuaciones

$$\begin{aligned} f_1(x) &= f_1(x_1, \dots, x_n) = 0, \\ &\vdots \\ f_m(x) &= f_m(x_1, \dots, x_n) = 0 \end{aligned}$$

es localmente resoluble para $x_1, \dots, x_m$. Esto significa lo siguiente: existen una vecindad abierta Y del punto $\xi^0 = (x_{m+1}^0, \dots, x_n^0) \in \mathbb{R}^{n-m}$ y m funciones unívocamente definidas $\varphi_\mu : \mathbb{R}^{n-m} \rightarrow \mathbb{R}$ ($\mu = 1, \dots, m$) de las variables $x_{m+1}, \dots, x_n$,

$$\begin{aligned} x_1 &= \varphi_1(x_{m+1}, \dots, x_n), \\ &\vdots \\ x_m &= \varphi_m(x_{m+1}, \dots, x_n), \end{aligned}$$

todas de ellas definidas sobre Y . Estas funciones satisfacen lo siguiente:

1. $\varphi_\mu \in C^1(Y)$ para $\mu = 1, \dots, m$.
2. Para todo $(x_{m+1}, \dots, x_n) \in Y$ se tiene que

$$\begin{aligned} f_1(\varphi_1(x_{m+1}, \dots, x_n), \dots, \varphi_m(x_{m+1}, \dots, x_n), x_{m+1}, \dots, x_n) &= 0, \\ &\vdots \\ f_m(\varphi_1(x_{m+1}, \dots, x_n), \dots, \varphi_m(x_{m+1}, \dots, x_n), x_{m+1}, \dots, x_n) &= 0. \end{aligned}$$

3. Las matrices

$$\phi(\xi^0) = \left(\frac{\partial \varphi_\mu}{\partial x_\nu}(\xi^0) \right)_{\substack{\mu=1, \dots, m \\ \nu=m+1, \dots, n}} \quad y \quad \tilde{F}(x^0) = \left(\frac{\partial f_\mu}{\partial x_\nu}(x^0) \right)_{\substack{\mu=1, \dots, m \\ \nu=m+1, \dots, n}}$$

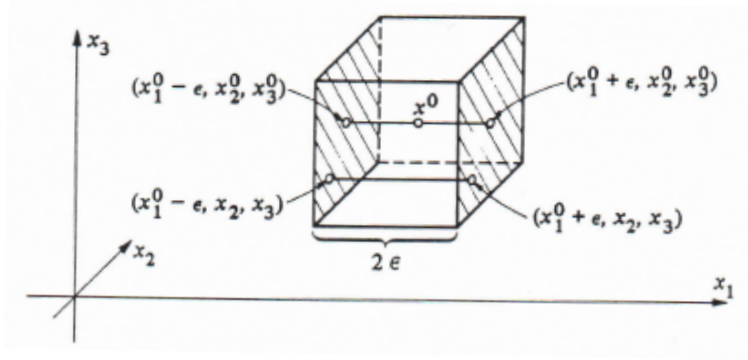


FIGURA 3.1. Ilustración de la demostración del Teorema 3.2.

satisfacen la relación

$$\phi(\xi^0) = -(F(x^0))^{-1} \cdot \tilde{F}(x^0).$$

Demostración. La demostración se realizará por inducción sobre μ , es decir, el número de ecuaciones. Primeramente notamos que después de posiblemente reenumerar las funciones f_μ y las m primeras variables $x_1, \dots, x_m$ podemos suponer que todos los subdeterminantes principales de (3.3) satisfacen

$$\begin{vmatrix} \frac{\partial f_1}{\partial x_1}(x^0) & \dots & \frac{\partial f_1}{\partial x_\mu}(x^0) \\ \vdots & & \vdots \\ \frac{\partial f_\mu}{\partial x_1}(x^0) & \dots & \frac{\partial f_\mu}{\partial x_\mu}(x^0) \end{vmatrix} \neq 0, \quad \mu = 1, \dots, m.$$

1. Sea $\mu = 1$. Según hipótesis,

$$\frac{\partial f_1}{\partial x_1}(x^0) \neq 0;$$

sin pérdida de la generalidad podemos suponer que

$$\frac{\partial f_1}{\partial x_1}(x^0) > 0.$$

a) En virtud de la hipótesis $f_1 \in C^1(X)$ existe un $\varepsilon > 0$ tal que

$$f_1(x_1^0 - \varepsilon, x_2^0, \dots, x_n^0) < 0, \quad f_1(x_1^0 + \varepsilon, x_2^0, \dots, x_n^0) > 0,$$

además existe un $\delta = \delta_\varepsilon > 0$ apropiado tal que para cada selección de x_ν con $|x_\nu - x_\nu^0| < \delta$ ($\nu = 2, \dots, n$) se tiene que

$$f_1(x_1^0 - \varepsilon, x_2, \dots, x_n) < 0, \quad f_1(x_1^0 + \varepsilon, x_2, \dots, x_n) > 0.$$

Eligiendo ε y δ aún más pequeños si fuera necesario, podemos suponer que en el entero intervalo n -dimensional

$$|x_1 - x_1^0| < \varepsilon, \quad |x_\nu - x_\nu^0| < \delta \quad (\nu = 2, \dots, n)$$

se tiene que

$$\frac{\partial f_1}{\partial x_1}(x) > 0.$$

Ahora sea elegido $(x_2, \dots, x_n)$ tal que $|x_\nu - x_\nu^0| < \delta$ para $\nu = 2, \dots, n$. Entonces, dado que $f_1(x_1, x_2, \dots, x_n)$ es una función monótona creciente de x_1 en el intervalo $|x_1 - x_1^0| < \varepsilon$, existe exactamente un valor $x_1 = \varphi_1(x_2, \dots, x_n)$ tal que

$$f_1(\varphi_1(x_2, \dots, x_n), x_2, \dots, x_n) = 0.$$

Esto define sobre $|x_\nu - x_\nu^0| < \delta$ ($\nu = 2, \dots, n$) una función $\varphi_1(x_2, \dots, x_n)$. Según la construcción, siempre se tiene que

$$|x_1 - x_1^0| = |\varphi_1(x_2, \dots, x_n) - \varphi_1(x_2^0, \dots, x_n^0)| < \varepsilon.$$

Puesto que la misma demostración puede ser realizada para un $\varepsilon > 0$ arbitrariamente pequeño, esto implica la continuidad de φ_1 en $(x_2^0, \dots, x_n^0)$. Para cualquier otro punto $(x_2, \dots, x_n)$ tal que $|x_\nu - x_\nu^0| < \delta$ para $\nu = 2, \dots, n$ resulta la continuidad si repetimos la misma demostración remplazando x^0 por $(\varphi_1(x_2, \dots, x_n), x_2, \dots, x_n)$.

- b) El enunciado $\varphi_1 \in C^1$ es una consecuencia del Teorema del Valor Intermedio para f_1 aplicado a los puntos $(x_2^0, \dots, x_k^0, \dots, x_n^0)$ y $(x_2^0, \dots, x_k^0 + h, \dots, x_n^0)$, donde suponemos $|h| < \delta$ para asegurar que también el segundo punto pertenezca al dominio de φ_1 . Ahora, si escribimos detalladamente en el argumento $(x_2, \dots, x_n)$ solamente la k -ésima componente y definimos $x_1^h := \varphi_1(\dots, x_k^0 + h, \dots)$, se tiene con un $\vartheta \in (0, 1)$ que

$$\begin{aligned} 0 &= f_1(x_1^h, \dots, x_k^0 + h, \dots) - f_1(x_1^0, \dots, x_k^0, \dots) \\ &= \frac{\partial f_1}{\partial x_1}(x_1^0 + \vartheta(x_1^h - x_1^0), \dots, x_k^0 + \vartheta h, \dots)(x_1^h - x_1^0) \\ &\quad + \frac{\partial f_1}{\partial x_k}(x_1^0 + \vartheta(x_1^h - x_1^0), \dots, x_k^0 + \vartheta h, \dots)h. \end{aligned}$$

Después de dividir por h y tomando el límite $h \rightarrow 0$ obtenemos

$$\begin{aligned} 0 &= \frac{\partial f_1}{\partial x_1}(x_1^0, \dots, x_k^0, \dots, x_n^0) \frac{\partial \varphi_1}{\partial x_k}(x_2^0, \dots, x_k^0, \dots, x_n^0) \\ &\quad + \frac{\partial f_1}{\partial x_k}(x_1^0, \dots, x_k^0, \dots, x_n^0), \end{aligned}$$

Esto implica la existencia de la derivada parcial de φ_1 con respecto a x_k ($k = 2, \dots, n$) en el punto $(x_2^0, \dots, x_n^0)$. Para cualquier otro punto podemos proceder como en (a). Entonces el enunciado $\varphi_1 \in C^1$ es una consecuencia de la suposición $f_1 \in C^1$.

2. Supongamos ahora que el teorema haya sido demostrado para $\mu - 1$, donde $2 \leq \mu \leq m$. Según la presuposición sabemos que

$$\begin{vmatrix} \frac{\partial f_1}{\partial x_1}(x^0) & \dots & \frac{\partial f_1}{\partial x_{\mu-1}}(x^0) \\ \vdots & & \vdots \\ \frac{\partial f_{\mu-1}}{\partial x_1}(x^0) & \dots & \frac{\partial f_{\mu-1}}{\partial x_{\mu-1}}(x^0) \end{vmatrix} \neq 0.$$

Según la hipótesis de inducción las primeras $\mu - 1$ ecuaciones del sistema son localmente resolubles, es decir existen un $\tilde{\delta} > 0$ y $\mu - 1$ funciones

$$\begin{aligned} x_1 &= \tilde{\varphi}_1(x_\mu, \dots, x_n), \\ &\vdots \\ x_{\mu-1} &= \tilde{\varphi}_{\mu-1}(x_\mu, \dots, x_n) \end{aligned}$$

las que son definidas, continuas y pertenecen a C^1 para $|x_\nu - x_\nu^0| < \tilde{\delta}$ para $\nu = \mu, \dots, n$, de modo que allí se tiene que

$$\begin{aligned} f_1(\tilde{\varphi}_1(x_\mu, \dots, x_n), \dots, \tilde{\varphi}_{\mu-1}(x_\mu, \dots, x_n), x_\mu, \dots, x_n) &= 0, \\ &\vdots \\ f_{\mu-1}(\tilde{\varphi}_1(x_\mu, \dots, x_n), \dots, \tilde{\varphi}_{\mu-1}(x_\mu, \dots, x_n), x_\mu, \dots, x_n) &= 0. \end{aligned}$$

Tomando la derivada parcial de cada una de estas ecuaciones con respecto a x_μ obtenemos para todo punto $(x_\mu, \dots, x_n)$ con $|x_\nu - x_\nu^0| < \tilde{\delta}$ para $\nu = \mu, \dots, n$ las siguientes identidades:

$$\begin{aligned} \frac{\partial f_1}{\partial x_1} \frac{\partial \tilde{\varphi}_1}{\partial x_\mu} + \dots + \frac{\partial f_1}{\partial x_{\mu-1}} \frac{\partial \tilde{\varphi}_{\mu-1}}{\partial x_\mu} + \frac{\partial f_1}{\partial x_\mu} &= 0, \\ &\vdots \\ \frac{\partial f_{\mu-1}}{\partial x_1} \frac{\partial \tilde{\varphi}_1}{\partial x_\mu} + \dots + \frac{\partial f_{\mu-1}}{\partial x_{\mu-1}} \frac{\partial \tilde{\varphi}_{\mu-1}}{\partial x_\mu} + \frac{\partial f_{\mu-1}}{\partial x_\mu} &= 0. \end{aligned} \tag{3.4}$$

Consideremos ahora para $(x_\mu, \dots, x_n)$ con $|x_\nu - x_\nu^0| < \tilde{\delta}$ ($\nu = \mu, \dots, n$) la μ -ésima ecuación del sistema, donde reemplazamos $x_1, \dots, x_{\mu-1}$ por las funciones $\tilde{\varphi}_1, \dots, \tilde{\varphi}_{\mu-1}$:

$$\tilde{f}_\mu(x_\mu, \dots, x_n) = f_\mu(\tilde{\varphi}_1(x_\mu, \dots, x_n), \dots, \tilde{\varphi}_{\mu-1}(x_\mu, \dots, x_n), x_\mu, \dots, x_n) = 0.$$

Por supuesto, esta ecuación no está automáticamente satisfecha para puntos arbitrarios $(x_\mu, \dots, x_n)$ con $|x_\nu - x_\nu^0| < \tilde{\delta}$ ($\nu = \mu, \dots, n$); más bien hay que despejar x_μ de esta ecuación.

Para tal efecto, y para poder aplicar el paso (1) de esta demostración, debemos demostrar primeramente que

$$\frac{\partial \tilde{f}_\mu}{\partial x_\mu}(x_\mu^0, \dots, x_n^0) \neq 0. \tag{3.5}$$

Pero para $(x_\mu, \dots, x_n)$ con $|x_\nu - x_\nu^0| < \tilde{\delta}$ ($\nu = \mu, \dots, n$) se tiene que

$$\frac{\partial \tilde{f}_\mu}{\partial x_\mu} = \frac{\partial f_\mu}{\partial x_1} \frac{\partial \tilde{\varphi}_1}{\partial x_\mu} + \dots + \frac{\partial f_\mu}{\partial x_{\mu-1}} \frac{\partial \tilde{\varphi}_{\mu-1}}{\partial x_\mu} + \frac{\partial f_\mu}{\partial x_\mu}.$$

Si esta expresión desapareciera al insertar $(x_\mu^0, \dots, x_n^0)$, tendríamos junto con las ecuaciones (3.4) un sistema de ecuaciones homogéneo con un determinante no nulo pero con la solución no trivial

$$\left(\frac{\partial \tilde{\varphi}_1}{\partial x_\mu}(x_\mu^0, \dots, x_n^0), \dots, \frac{\partial \tilde{\varphi}_{\mu-1}}{\partial x_\mu}(x_\mu^0, \dots, x_n^0), 1 \right),$$

una contradicción. Entonces (3.5) es válido, y existen un $\delta \in (0, \tilde{\delta})$ y una función

$$x_\mu = \varphi_\mu(x_{\mu+1}, \dots, x_n) \quad (3.6)$$

la cual está definida, continua, y pertenece a C^1 para $(x_{\mu+1}, \dots, x_n)$ con $|x_\nu - x_\nu^0| < \tilde{\delta}$ para $\nu = \mu + 1, \dots, n$ tal que

$$\begin{aligned} f_\mu \left(\tilde{\varphi}_1(\varphi_\mu(x_{\mu+1}, \dots, x_n), x_{\mu+1}, \dots, x_n), \dots, \right. \\ \left. \tilde{\varphi}_{\mu-1}(\varphi_\mu(x_{\mu+1}, \dots, x_n), x_{\mu+1}, \dots, x_n), \varphi_\mu(x_{\mu+1}, \dots, x_n), x_{\mu+1}, \dots, x_n) \right) = 0. \end{aligned}$$

Si todavía definimos

$$\begin{aligned} x_1 &= \tilde{\varphi}_1(\varphi_\mu(x_{\mu+1}, \dots, x_n), x_{\mu+1}, \dots, x_n) =: \varphi_1(x_{\mu+1}, \dots, x_n), \\ &\vdots \\ x_{\mu-1} &= \tilde{\varphi}_{\mu-1}(\varphi_\mu(x_{\mu+1}, \dots, x_n), x_{\mu+1}, \dots, x_n) =: \varphi_{\mu-1}(x_{\mu+1}, \dots, x_n), \end{aligned}$$

hemos encontrado junto con (3.6) las soluciones buscadas.

La afirmación $\varphi_1, \dots, \varphi_\mu \in C^1$ sigue de las propiedades correspondientes sobre las funciones φ_μ y $\tilde{\varphi}_1, \dots, \tilde{\varphi}_{\mu-1}$ mediante la aplicación de la regla de la cadena. Asimismo, la afirmación (3) del teorema es una consecuencia de la regla de la cadena. ■

Ejemplo 3.4. Consideremos la ecuación

$$f(x, y, z) = y^2 + xz + z^2 - e^z - 4 = 0$$

y queremos resolverla para z en una vecindad de $(0, e, 2)$. Para la matriz F obtenemos aquí

$$\det F = \frac{\partial f}{\partial z}(0, e, 2) = (x + 2z - e^z)|_{(0, e, 2)} = 4 - e^2 \neq 0.$$

Según el Teorema 3.2, esta ecuación puede ser resuelta en una vecindad de $(0, e)$ en la forma $z = z(x, y)$ de tal manera que $z(0, e) = 2$.

Queremos calcular las derivadas $\partial z / \partial x(0, e)$ y $\partial z / \partial y(0, e)$. Aplicando la regla de la cadena a $f(x, y, z) = 0$ obtenemos por "derivación implícita" con respecto a x e y las ecuaciones

$$0 = z + x \frac{\partial z}{\partial x} + 2z \frac{\partial z}{\partial x} - e^z \frac{\partial z}{\partial x}, \quad 0 = 2y + x \frac{\partial z}{\partial y} + 2z \frac{\partial z}{\partial y} - e^z \frac{\partial z}{\partial y},$$

es decir,

$$\frac{\partial z}{\partial x} = -\frac{z}{x + 2z - e^z}, \quad \frac{\partial z}{\partial y} = -\frac{2y}{x + 2z - e^z}.$$

Insertando $x = 0$, $y = e$ y $z = z(0, e) = 2$ obtenemos

$$\frac{\partial z}{\partial x}(0, e) = -\frac{2}{4 - e^2}, \quad \frac{\partial z}{\partial y}(0, e) = -\frac{2e}{4 - e^2}.$$

Note que ha sido posible calcular estas derivadas parciales sin conocer explícitamente la función $z(x, y)$; solamente necesitamos conocer el valor de $z(x, y)$ en el punto $(0, e)$.

Ejemplo 3.5. Consideremos el sistema de ecuaciones

$$\begin{aligned} f_1(x_1, x_2, x_3, x_4) &= x_1x_3 + x_2x_4^2 = 0, \\ f_2(x_1, x_2, x_3, x_4) &= x_1x_4^3 + x_2^2x_3^6 = 0. \end{aligned}$$

Evidentemente el sistema está satisfecho en el punto

$$(x_1^0, x_2^0, x_3^0, x_4^0) = (0, 1, 0, 0).$$

Queremos analizar si podemos resolver el sistema para $x_3 = x_3(x_1, x_2)$, $x_4 = x_4(x_1, x_2)$. Aquí obtenemos

$$F = \begin{bmatrix} \frac{\partial f_1}{\partial x_3} & \frac{\partial f_1}{\partial x_4} \\ \frac{\partial f_2}{\partial x_3} & \frac{\partial f_2}{\partial x_4} \end{bmatrix} = \begin{bmatrix} x_1 & 2x_2x_4 \\ 6x_2^2x_3^5 & 3x_1x_4^2 \end{bmatrix}.$$

Para el punto considerado,

$$\det F = 3x_1^2x_4^2 - 12x_2^3x_3^5x_4 = 0.$$

En este caso, el Teorema Principal de Funciones Implícitas (Teorema 3.2) no nos entrega ninguna información acerca de la existencia y unicidad de una solución del sistema en una vecindad del punto considerado.

3.3. Funciones inversas

El siguiente teorema representa una de las aplicaciones más importantes del Teorema Principal de Funciones Implícitas.

Teorema 3.3 (Teorema Principal de las Funciones Inversas). *Se considera una aplicación $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$. Sea $X \subset D(f)$ un conjunto abierto, y sea $f \in C^1(X)$. En el punto x^0 el Jacobiano de f sea regular, es decir,*

$$\det \left(\frac{df}{dx}(x^0) \right) \neq 0.$$

1. *En este caso existen una vecindad abierta $U(x^0)$ de x^0 y una vecindad abierta $V(y^0)$ de $y^0 = f(x^0)$ tales que $U(x^0)$ es mapeado de manera biyectiva a $V(y^0)$, es decir sobre $V(y^0)$ existe la aplicación inversa f^{-1} .*

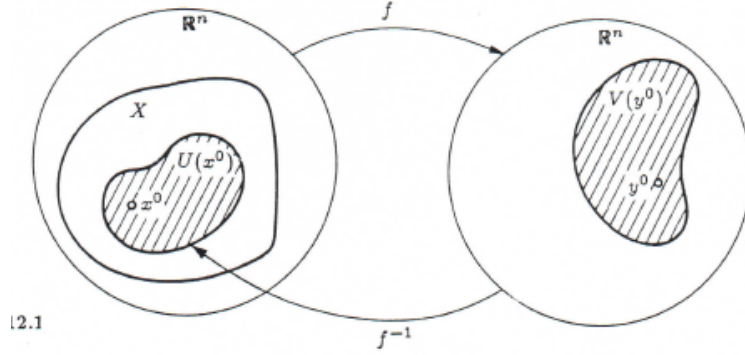


FIGURA 3.2. Ilustración del Teorema 3.3.

2. Esta aplicación inversa satisface $f^{-1} \in C^1(V(y^0))$, y para $y = f(x) \in V(y^0)$ se tiene la siguiente relación entre los Jacobianos df/dx y df^{-1}/dy :

$$\frac{df^{-1}}{dy}(y) = \left(\frac{df}{dx}(x) \right)^{-1}.$$

Demostración. Escribimos la aplicación dada

$$\begin{aligned} y_1 &= f_1(x_1, \dots, x_n), \\ &\vdots \\ y_n &= f_n(x_1, \dots, x_n) \end{aligned}$$

como un sistema de ecuaciones en $\mathbb{R}^{2n}$, donde ponemos $y_1 = x_{n+1}, \dots, y_n = x_{2n}$ y $y_1^0 = x_{n+1}^0, \dots, y_n^0 = x_{2n}^0$:

$$\begin{aligned} F_1(x_1, \dots, x_{2n}) &= f_1(x_1, \dots, x_n) - x_{n+1} = 0, \\ &\vdots \\ F_n(x_1, \dots, x_{2n}) &= f_n(x_1, \dots, x_n) - x_{2n} = 0. \end{aligned} \tag{3.7}$$

En el punto $(x_1^0, \dots, x_{2n}^0)$ el sistema de ecuaciones está satisfecho; además, en una vecindad de este punto $F = (F_1, \dots, F_n)$ pertenece a C^1 , y se tiene que

$$\begin{vmatrix} \frac{\partial F_1}{\partial x_1} & \dots & \frac{\partial F_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial F_n}{\partial x_1} & \dots & \frac{\partial F_n}{\partial x_n} \end{vmatrix} (x_1^0, \dots, x_{2n}^0) = \begin{vmatrix} \frac{\partial f_1}{\partial x_1} & \dots & \frac{\partial f_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial f_n}{\partial x_1} & \dots & \frac{\partial f_n}{\partial x_n} \end{vmatrix} (x_1^0, \dots, x_n^0) \neq 0.$$

Esto significa que localmente podemos despejar $x_1, \dots, x_n$, es decir existen un $\delta > 0$ y funciones

$$\begin{aligned} x_1 &= f_1^{-1}(x_{n+1}, \dots, x_{2n}), \\ &\vdots \\ x_n &= f_n^{-1}(x_{n+1}, \dots, x_{2n}) \end{aligned}$$

definidas, continuas y C^1 sobre $|x_\nu - x_\nu^0| < \delta$ ($\nu = n+1, \dots, 2n$), las cuales allí satisfacen (3.7). Si nuevamente definimos $y_\nu = x_{n+\nu}$ y $y_\nu^0 = x_{n+\nu}^0$ para $\nu = 1, \dots, n$ esto significa que para $(y_1, \dots, y_n)$ tales que $|y_\nu - y_\nu^0| < \delta$ ($\nu = 1, \dots, n$) hemos encontrado funciones tales que

$$\begin{aligned} y_1 &= f_1(f_1^{-1}(y_1, \dots, y_n), \dots, f_n^{-1}(y_1, \dots, y_n)), \\ &\vdots \\ y_n &= f_n(f_1^{-1}(y_1, \dots, y_n), \dots, f_n^{-1}(y_1, \dots, y_n)). \end{aligned}$$

La derivación parcial nos entrega

$$\frac{\partial y_\mu}{\partial y_\nu} = \sum_{i=1}^n \frac{\partial f_\mu}{\partial x_i} \frac{\partial f_i^{-1}}{\partial y_\nu} = \begin{cases} 1 & \text{si } \mu = \nu, \\ 0 & \text{si } \mu \neq \nu. \end{cases}$$

Esto demuestra el enunciado (2) según la definición del producto matricial. ■

Ejemplo 3.6. Consideremos la aplicación $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ definida por

$$y_1 = f_1(x_1, x_2) = x_1^2, \quad y_2 = f_2(x_1, x_2) = x_1 + x_2.$$

Aquí

$$\left| \frac{df}{dx} \right| = \begin{vmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} \end{vmatrix} = \begin{vmatrix} 2x_1 & 0 \\ 1 & 1 \end{vmatrix} = 2x_1.$$

Consideremos un punto $x^0 = (x_1^0, x_2^0) \in \mathbb{R}^2$, entonces

$$\left| \frac{df}{dx}(x^0) \right| \neq 0$$

si y sólo si $x_1^0 \neq 0$, es decir si x^0 está localizado o en el semiplano izquierdo o en el semiplano derecho.

Supongamos primero que $x_1^0 > 0$. En este caso la aplicación inversa f_R^{-1} de f está definida por

$$x_1 = \sqrt{y_1}, \quad x_2 = y_2 - \sqrt{y_1}.$$

El dominio de f_R^{-1} es el semiplano derecho; la imagen igualmente es el semiplano derecho.

Para $x_1^0 < 0$, obtenemos la función inversa f_L^{-1} de f definida por

$$x_1 = -\sqrt{y_1}, \quad x_2 = y_2 + \sqrt{y_1}.$$

Nuevamente, el dominio de f_L^{-1} es el semiplano derecho; la imagen ahora es el semiplano izquierdo.

Ejemplo 3.7. Sea la función $f := (u, v) : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ dada por

$$u = u(x, y) = x \cos y, \quad v = v(x, y) = x \sin y,$$

sobre $D(f) = \{(x, y) \in \mathbb{R}^2 \mid x > 0\}$. Aquí queremos analizar la invertibilidad de f . Para tal efecto notamos que aquí el Jacobiano está dado por

$$\frac{\partial(u, v)}{\partial(x, y)} = \begin{bmatrix} \cos y & -x \sin y \\ \sin y & x \cos y \end{bmatrix}$$

con el determinante

$$\det \frac{\partial(u, v)}{\partial(x, y)} = x \cos^2 y + x \sin^2 y = x.$$

Según el Teorema 3.3 podemos concluir que f es invertible en una vecindad de cada punto (x_0, y_0) con $x_0 \neq 0$. Pero f no es invertible globalmente, dado que

$$\begin{aligned} u(x_0, y_0 + 2\pi) &= x_0 \cos(y_0 + 2\pi) = x_0 \cos y_0 = u(x_0, y_0), \\ v(x_0, y_0 + 2\pi) &= x_0 \sin(y_0 + 2\pi) = x_0 \sin y_0 = v(x_0, y_0), \end{aligned}$$

así que la imagen de los puntos (x_0, y_0) y $(x_0, y_0 + 2\pi)$ siempre es la misma.

Ejemplo 3.8 (Problema 4, Certamen 1, Curso 2009/I). Se consideran las funciones

$$\begin{aligned} f_1(x, y, u, v) &= x + y + (u + v)(u - v), \\ f_2(x, y, u, v) &= 2x - y + (u^2 - 1) \cos v + 1. \end{aligned} \quad (3.8)$$

Sea $P = (x_0 = 0, y_0 = 0, u_0 = 0, v_0 = 0)$.

- a) Analizar si en una vecindad de P , existen funciones $g_1 = g_1(u, v)$ y $g_2 = g_2(u, v)$ tales que las variables $x = g_1(u, v)$ e $y = g_2(u, v)$ pueden ser despejadas de

$$\begin{aligned} f_1(x, y, u, v) &= 0, \\ f_2(x, y, u, v) &= 0. \end{aligned} \quad (3.9)$$

Si posible, calcular las derivadas parciales

$$\frac{\partial g_1}{\partial u}(0, 0), \quad \frac{\partial g_1}{\partial v}(0, 0), \quad \frac{\partial g_2}{\partial u}(0, 0), \quad \frac{\partial g_2}{\partial v}(0, 0).$$

- b) Analizar si en una vecindad de P , existen funciones $h_1 = h_1(x, y)$ y $h_2 = h_2(x, y)$ tales que las variables $u = h_1(x, y)$ y $v = h_2(x, y)$ pueden ser despejadas de (3.9). Si posible, calcular las derivadas parciales

$$\frac{\partial h_1}{\partial x}(0, 0), \quad \frac{\partial h_1}{\partial y}(0, 0), \quad \frac{\partial h_2}{\partial x}(0, 0), \quad \frac{\partial h_2}{\partial y}(0, 0).$$

Solución sugerida.

- a) Aquí calculamos

$$\det \begin{bmatrix} \frac{\partial f_1}{\partial x} & \frac{\partial f_1}{\partial y} \\ \frac{\partial f_2}{\partial x} & \frac{\partial f_2}{\partial y} \end{bmatrix} = \begin{vmatrix} 1 & 1 \\ 2 & -1 \end{vmatrix} = -3 \neq 0,$$

entonces, según el Teorema de las Funciones Implícitas, las funciones deseadas g_1 y g_2 existen. Por derivación implícita obtenemos

$$\begin{aligned} 0 &= \frac{\partial g_1}{\partial u} + \frac{\partial g_2}{\partial u} + 2u, & 0 &= 2\frac{\partial g_1}{\partial u} - \frac{\partial g_2}{\partial u} + 2u \cos v, \\ 0 &= \frac{\partial g_1}{\partial v} + \frac{\partial g_2}{\partial v} - 2v, & 0 &= 2\frac{\partial g_1}{\partial v} - \frac{\partial g_2}{\partial v} - (u^2 - 1) \sin v, \end{aligned}$$

es decir en P_0 las derivadas parciales son soluciones de

$$\begin{bmatrix} 1 & 1 \\ 2 & -1 \end{bmatrix} \begin{pmatrix} (\partial g_1 / \partial u)(0, 0) \\ (\partial g_2 / \partial u)(0, 0) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad \begin{bmatrix} 1 & 1 \\ 2 & -1 \end{bmatrix} \begin{pmatrix} (\partial g_1 / \partial v)(0, 0) \\ (\partial g_2 / \partial v)(0, 0) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

de lo cual se desprende que

$$\frac{\partial g_1}{\partial u}(0, 0) = \frac{\partial g_1}{\partial v}(0, 0) = \frac{\partial g_2}{\partial u}(0, 0) = \frac{\partial g_2}{\partial v}(0, 0) = 0.$$

b) Aquí calculamos para $u = v = 0$

$$\det \begin{bmatrix} \frac{\partial f_1}{\partial u} & \frac{\partial f_1}{\partial v} \\ \frac{\partial f_2}{\partial u} & \frac{\partial f_2}{\partial v} \end{bmatrix} = \begin{vmatrix} 2u & -2v \\ 2u \cos v & (1 - u^2) \sin v \end{vmatrix} = 0,$$

entonces el Teorema de las Funciones Implícitas no puede ser aplicado. Efectivamente, las variables u y v no pueden ser despejadas de manera única en ninguna vecindad de P . Esto es debido a que las funciones f_1 y f_2 son simétricas con respecto a u y v :

$$f_i(x, y, u, v) = f_i(x, y, \pm u, \pm v), \quad i = 1, 2.$$

Dado que $(u, v) = (0, 0)$ es una solución de (3.9) solamente para $(x = 0, y = 0)$, vemos que para cada $(x, y) \neq (0, 0)$, la solución de (3.8) entrega $(u, v) \neq (0, 0)$, es decir considerando todas las combinaciones de signos en $(\pm u, \pm v)$, existe por lo menos una segunda solución.

Finalmente notamos el siguiente resultado como aplicación del Teorema Principal de las Funciones Inversas.

Teorema 3.4. Sea $G \subset \mathbb{R}^n$ una región. La aplicación $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ satisfaga $f \in C^1(G)$, y sea

$$\det \frac{df}{dx}(x) \neq 0 \quad \text{para todo } x \in G. \quad (3.10)$$

Entonces también $f(G)$ es una región.

Demostración. Según el Teorema 2.18, el conjunto G es abierto y conexo. Puesto que f es continua sobre G , el Teorema 1.18 implica que $f(G)$ es conexo. Sean $y^0 \in f(G)$ y $x^0 \in G$ escogidos tales que $f(x^0) = y^0$. Entonces, en virtud de

$$\det \frac{df}{dx}(x^0) \neq 0,$$

según el Teorema 3.3 existe una vecindad $V(y^0)$ de y^0 con $V(y^0) \subset f(G)$. Puesto que $y^0 \in f(G)$ es arbitrario, $f(G)$ es abierto y según el Teorema 2.18 una región. ■

Comentamos que el enunciado del Teorema 3.4 ya no es válido si la presuposición (3.10) no es válida. Por ejemplo, consideremos la región $G = (-2, 2) \subset \mathbb{R}$ y la función $f(x) = \sin x$. Se tiene que $f \in C^1(G)$, pero $f(G) = [-1, 1]$ no es una región.

3.4. Extremos de funciones de varias variables

Teorema 3.5. *Sea la función $f : \mathbb{R}^n \rightarrow \mathbb{R}$ parcialmente diferenciable con respecto a cada variable en un punto interior x^0 de $D(f)$. Si f posee un extremo relativo en x^0 , entonces se tiene que*

$$\nabla f(x^0) = \vec{0}.$$

Demostración. El enunciado sigue inmediatamente del hecho que f posee en x^0 un extremo en cada dirección de coordenada $\vec{e}_i$ como función de la variable escalar x_i , lo que implica $f_{x_i}(x^0) = 0$ para $i = 1, \dots, n$. ■

Comentamos que la condición indicada en el Teorema 3.5, $\nabla f(x^0) = \vec{0}$, es una condición necesaria, pero no suficiente para la existencia de un extremo relativo. Para ilustrar eso, consideremos la función $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ con $D(f) = \mathbb{R}^2$ y $f(x, y) = xy$ (ver Figura 2.1). Aquí obtenemos las derivadas parciales $f_x(x, y) = y$ y $f_y(x, y) = x$, por lo tanto $\nabla f(x, y) = \vec{0}$ si y sólo si $(x_0, y_0) = (0, 0)$. Sin embargo, la función f no posee un extremo relativo en $(0, 0)$ dado que en cada vecindad de $(0, 0)$ existen valores de f positivo y negativos.

Aquellos puntos x^0 de una función $f : \mathbb{R}^n \rightarrow \mathbb{R}$ donde se tiene que $\nabla f(x^0) = \vec{0}$ pero que no son extremos relativos se llaman *puntos de silla*.

Para la formulación de condiciones suficientes para la existencia de extremos locales de funciones $f : \mathbb{R}^n \rightarrow \mathbb{R}$ necesitamos estudiar formas cuadráticas.

Definición 3.3. *Sea $A = (a_{ik})_{i,k=1,\dots,n}$ una matriz simétrica.*

1. *El polinomio $Q_A : \mathbb{R}^n \rightarrow \mathbb{R}$ definido por*

$$Q_A(x) = \sum_{i,k=1}^n a_{ik}x_i x_k, \quad x = (x_1, \dots, x_n) \in \mathbb{R}^n$$

se llama la forma cuadrática asociada con A .

2. *La matriz A o la función Q_A se llama semidefinida positiva si $Q_A(x) \geq 0$ para todo $x \in \mathbb{R}^n$ y semidefinida negativa si $Q_A(x) \leq 0$ para todo $x \in \mathbb{R}^n$.*
3. *La matriz A o la función Q_A se llama definida positiva si $Q_A(x) > 0$ para todo $x \in \mathbb{R}^n$, $x \neq (0, \dots, 0)$, y definida negativa si $Q_A(x) < 0$ para todo $x \in \mathbb{R}^n$, $x \neq (0, \dots, 0)$.*
4. *La matriz A o la función Q_A se llama indefinida si no es semidefinida ni positiva ni negativa.*

Teorema 3.6. *Sea $A = (a_{ik}) \in \mathbb{R}^{n \times n}$ una matriz simétrica; y sea $-A = (-a_{ik})$.*

1. *La matriz A es semidefinida positiva si y sólo si $-A$ es semidefinida negativa.*
2. *La matriz A es definida positiva si y sólo si $-A$ es definida negativa.*

Demostración. Las afirmaciones son una consecuencia inmediata de

$$Q_A(x) = \sum_{i,k=1}^n a_{ik}x_i x_k = - \sum_{i,k=1}^n (-a_{ik})x_i x_k = -Q_{(-A)}(x).$$

Teorema 3.7. *La matriz*

$$A = \begin{bmatrix} a & b \\ b & d \end{bmatrix}$$

es

- *semidefinida positiva si y sólo si $ad - b^2 \geq 0$, $a \geq 0$ y $d \geq 0$,*
- *semidefinida negativa si y sólo si $ad - b^2 \geq 0$, $a \leq 0$ y $d \leq 0$,*
- *definida positiva si y sólo si $ad - b^2 > 0$ y $a > 0$,*
- *definida negativa si y sólo si $ad - b^2 > 0$ y $a < 0$,*
- *indefinida si y sólo si $ad - b^2 < 0$.*

Demostración. Tarea. Se recomienda utilizar que para $(x, y) \in \mathbb{R}^2$ el polinomio $Q_A(x, y) = ax^2 + 2bxy + dy^2$ satisface

$$aQ_A(x, y) = (ax + by)^2 + (ad - b^2)y^2.$$

En general es muy difícil decidir si una matriz es definida; los métodos especializados son tópicos del álgebra lineal. Aquí mencionamos el siguiente criterio útil (sin demostración).

Teorema 3.8. *Se considera la matriz simétrica $A = (a_{ik})$ con las submatrices principales*

$$A_\nu = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1\nu} \\ a_{21} & a_{22} & \dots & a_{2\nu} \\ \vdots & \vdots & & \vdots \\ a_{\nu 1} & a_{\nu 2} & \dots & a_{\nu\nu} \end{bmatrix}.$$

Sea $\det A_\nu$ el determinante de A_ν . Entonces

1. *A es definida positiva si y sólo si*

$$\det A_\nu > 0 \quad \text{para } \nu = 1, \dots, n,$$

2. *A es definida negativa si y sólo si*

$$(-1)^\nu \det A_\nu > 0 \quad \text{para } \nu = 1, \dots, n.$$

Además, se tiene el siguiente resultado.

Teorema 3.9. *Para un $\varepsilon > 0$ sean*

$$U'_\varepsilon := \left\{ x \in \mathbb{R}^n \mid 0 < \sqrt{x_1^2 + \dots + x_n^2} < \varepsilon \right\},$$

$$S_\varepsilon := \left\{ x \in \mathbb{R}^n \mid d(x, 0) = \sqrt{x_1^2 + \dots + x_n^2} = \varepsilon \right\}.$$

Si sobre U'_ε o sobre S_ε se tiene que

- *$Q_A(x) \geq 0$, entonces A es semidefinida positiva,*
- *$Q_A(x) \leq 0$, entonces A es semidefinida negativa,*
- *$Q_A(x) > 0$, entonces A es definida positiva,*
- *$Q_A(x) < 0$, entonces A es definida negativa.*

Demostración. Demostraremos solamente los primeros dos enunciados; la demostración de los dos demás enunciados es análoga.

1. Sea $Q_A \geq 0$ sobre U'_ε . Entonces para $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ con $x \neq (0, \dots, 0)$ existe un $R > 0$ tal que

$$y = \left(\frac{x_1}{R}, \dots, \frac{x_n}{R} \right) \in U'_\varepsilon.$$

Ahora $Q_A(y) \geq 0$, y por lo tanto

$$Q_A(x) = R^2 Q_A(y) \geq 0.$$

2. Sobre S_ε sea $Q_A \geq 0$. Entonces para $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ con $x \neq (0, \dots, 0)$ sabemos que

$$y = \left(\frac{\varepsilon x_1}{d(x, 0)}, \dots, \frac{\varepsilon x_n}{d(x, 0)} \right) \in S_\varepsilon,$$

por lo tanto $Q_A(y) \geq 0$, y se tiene que

$$Q_A(x) = \frac{d^2(x, 0)}{\varepsilon^2} Q_A(y) \geq 0.$$

3. Para $Q_A \leq 0$ el enunciado es una consecuencia del Teorema 3.6.

■

Teorema 3.10. *Se considera la función $f : \mathbb{R}^n \rightarrow \mathbb{R}$. Sea $x^0 \in D(f)$, y para una vecindad $U(x^0)$ de x^0 sea $f \in C^2(U(x^0))$. Además, sea $\nabla f(x^0) = \vec{0}$, y consideremos la matriz*

$$A(x^0) = (f_{x_i x_k}(x^0))_{i,k=1,\dots,n}.$$

Entonces se tiene lo siguiente.

1. Si $A(x^0)$ es definida positiva, entonces f posee un mínimo local en x^0 . Si $A(x^0)$ es definida negativa, entonces f posee un máximo local en x^0 .
2. Si f posee un mínimo relativo en x^0 , entonces $A(x^0)$ es semidefinida positiva. Si f posee un máximo relativo en x^0 , entonces $A(x^0)$ es semidefinida negativa.
3. Si $A(x^0)$ es indefinida, entonces f no posee en x^0 ni un mínimo relativo ni un máximo relativo.

Demostración. Comentamos primeramente que según el Teorema 2.10, la matriz $A(x^0)$ es simétrica. Según el Teorema de Taylor existe un $\delta > 0$ tal que para todo $\vec{h} = \{h_1, \dots, h_n\}$ con $\|\vec{h}\| < \delta$ y un $\vartheta = \vartheta(\vec{h}) \in (0, 1)$ se tiene que

$$\begin{aligned} f(x^0 + \vec{h}) - f(x^0) &= (\vec{h} \cdot \nabla f)(x^0) + \frac{1}{2!} (\vec{h} \cdot \nabla f)^2(x^0 + \vartheta(\vec{h})\vec{h}) \\ &= \frac{1}{2} \sum_{i,k=1}^n f_{x_i x_k}(x^0 + \vartheta(\vec{h})\vec{h}) h_i h_k. \end{aligned} \tag{3.11}$$

1. Sea la matrix $A(x^0)$ definida positiva. Entonces la forma cuadrática

$$\sum_{i,k=1}^n f_{x_i x_k}(x^0) h_i h_k,$$

como función continua, asume un mínimo absoluto $m > 0$ sobre el conjunto compacto $S_1 = \{\vec{h} \mid \|\vec{h}\| = 1\}$, es decir sobre S_1 se tiene que

$$\sum_{i,k=1}^n f_{x_i x_k}(x^0) h_i h_k \geq m > 0.$$

Puesto que todas las funciones $f_{x_i x_k}(x)$ son continuas en x^0 , existe un $\delta' \in (0, \delta)$ tal que para todo x con $d(x, x^0) < \delta'$ y todo $\vec{h} \in S_1$ la desigualdad

$$\sum_{i,k=1}^n f_{x_i x_k}(x) h_i h_k > 0$$

es válida. Según el Teorema 3.9, entonces la forma cuadrática

$$\sum_{i,k=1}^n f_{x_i x_k}(x) h_i h_k$$

es definida positiva para todo x con $d(x, x^0) < \delta'$. Como consecuencia de (3.11), aplicado para todo $\vec{h}$ con $\|\vec{h}\| < \delta'$, obtenemos que

$$f(x^0 + \vec{h}) - f(x^0) > 0,$$

es decir f posee un mínimo relativo en x^0 .

Si $A(x^0)$ es definida negativa, entonces $-A(x^0)$ es definida positiva, por lo tanto $-f$ posee un mínimo local en x^0 , es decir, f posee un máximo local.

2. Si f posee un mínimo relativo en x^0 , entonces existe un $\delta'' \in (0, \delta)$ tal que

$$\sum_{i,k=1}^n f_{x_i x_k}(x^0 + \vartheta(\vec{h})\vec{h}) h_i h_k \geq 0$$

para todo $\vec{h}$ con $\|\vec{h}\| < \delta''$. Fijemos un tal $\vec{h}$, entonces se tiene para $\lambda \in (0, 1)$ arbitrario que

$$\sum_{i,k=1}^n f_{x_i x_k}(x^0 + \vartheta(\lambda\vec{h})\lambda\vec{h}) (\lambda h_i) (\lambda h_k) \geq 0.$$

Después de la división por λ^2 y tomando el límite cuando $\lambda \rightarrow 0$ obtenemos

$$\sum_{i,k=1}^n f_{x_i x_k}(x^0) h_i h_k \geq 0;$$

esto es válido para todo $\vec{h}$ con $\|\vec{h}\| < \delta''$. Según el Teorema 3.9, la matriz $A(x^0)$ es semidefinida positiva.

Por otro lado, si f posee un máximo relativo en x^0 , entonces $-f$ posee un mínimo relativo, por lo tanto $-A(x^0)$ es semidefinida positiva y $A(x^0)$ es semidefinida negativa.

3. Si $A(x^0)$ es indefinida podemos concluir de (1) y (2) que f no puede tener ni un máximo relativo ni un mínimo relativo. ■

Ejemplo 3.9. Consideremos la función $f : \mathbb{R}^3 \rightarrow \mathbb{R}$ con $D(f) = \mathbb{R}^3$ y

$$f(x, y, z) = 35 - 6x + 2z + x^2 - 2xy + 2y^2 + 2yz + 3z^2.$$

Aquí obtenemos las derivadas parciales

$$f_x = -6 + 2x - 2y, \quad f_y = -2x + 4y + 2z, \quad f_z = 2 + 2y + 6z.$$

Esto implica que $(x_0, y_0, z_0) = (8, 5, -2)$ es el único punto donde $\nabla f = \vec{0}$. En este punto se tiene que

$$f_{xx} = 2, \quad f_{xy} = -2, \quad f_{yy} = 4, \quad f_{xz} = 0, \quad f_{yz} = 2, \quad f_{zz} = 6,$$

es decir, la matriz $A(x_0, y_0, z_0)$ está dada por

$$A(x_0, y_0, z_0) = \begin{bmatrix} 2 & -2 & 0 \\ -2 & 4 & 2 \\ 0 & 2 & 6 \end{bmatrix}.$$

Los determinantes de las submatrices principales son

$$|2| = 2, \quad \begin{vmatrix} 2 & -2 \\ -2 & 4 \end{vmatrix} = 4, \quad \begin{vmatrix} 2 & -2 & 0 \\ -2 & 4 & 2 \\ 0 & 2 & 6 \end{vmatrix} = 16.$$

Todos estos determinantes son positivos. Concluimos que según el Teorema 3.8 la matriz $A(x_0, y_0, z_0)$ es definida positiva, por lo tanto la función f posee en $(x_0, y_0, z_0) = (8, 5, -2)$ un mínimo local.

Consideremos ahora el caso particular de una función $f : \mathbb{R}^2 \rightarrow \mathbb{R}$.

Teorema 3.11. Se considera la función $f : \mathbb{R}^2 \rightarrow \mathbb{R}$. Sea $(x_0, y_0) \in D(f)$, y en para una vecindad $U(x_0, y_0)$ de (x_0, y_0) sea $f \in C^2(U(x_0, y_0))$. Sea $\nabla f(x_0, y_0) = \vec{0}$, y el discriminante $\delta(x_0, y_0)$ definido por

$$\delta(x_0, y_0) := f_{xx}(x_0, y_0)f_{yy}(x_0, y_0) - (f_{xy}(x_0, y_0))^2.$$

1. Si $\delta(x_0, y_0) > 0$, entonces f posee un extremo en (x_0, y_0) . Se trata de un máximo relativo si $f_{xx}(x_0, y_0) < 0$ y de un mínimo relativo si $f_{xx}(x_0, y_0) > 0$.
2. Su $\delta(x_0, y_0) < 0$, entonces f no posee ningún extremo en (x_0, y_0) .

Demostración. Consideremos la matriz simétrica

$$A = \begin{bmatrix} f_{xx}(x_0, y_0) & f_{xy}(x_0, y_0) \\ f_{yx}(x_0, y_0) & f_{yy}(x_0, y_0) \end{bmatrix}.$$

Utilizando el Teorema 3.7 obtenemos lo siguiente.

1. Si $\delta(x_0, y_0) > 0$ y $f_{xx}(x_0, y_0) < 0$, entonces A es definida negativa, y según el Teorema 3.10 la función f posee un máximo local en (x_0, y_0) .
2. Si $\delta(x_0, y_0) > 0$ y $f_{xx}(x_0, y_0) > 0$, entonces A es definida positiva, y según el Teorema 3.10 la función f posee un mínimo local en (x_0, y_0) .
3. Si $\delta(x_0, y_0) < 0$, entonces A es indefinida, y según el Teorema 3.10 la función f no posee ni un máximo local ni un mínimo local en (x_0, y_0) .

■

Ejemplo 3.10. Consideremos la función $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ dada por

$$f(x, y) = xy + x - y + 1, \quad D(f) = \mathbb{R}^2.$$

Aquí obtenemos las derivadas parciales

$$f_x(x, y) = y + 1, \quad f_y(x, y) = x - 1, \quad f_{xy}(x, y) = 1, \quad f_{xx}(x, y) = f_{yy}(x, y) = 0.$$

Los únicos candidatos a ser extremos de f son aquellos puntos (x_0, y_0) donde se tiene que $f_x(x_0, y_0) = f_y(x_0, y_0) = 0$; en este caso el único punto con esta propiedad es $(1, -1)$. Pero

$$\delta(1, -1) = f_{xx}(1, -1)f_{yy}(1, -1) - (f_{xy}(1, -1))^2 = -1 < 0,$$

entonces según el Teorema 3.11, f no posee ningún extremo.

Ejemplo 3.11. Consideremos la función $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ dada por

$$f(x, y) = x^3 - 12xy + 8y^3, \quad D(f) = \mathbb{R}^2.$$

Aquí obtenemos las derivadas parciales

$$\begin{aligned} f_x(x, y) &= 3x^2 - 12y, & f_y(x, y) &= -12x + 24y^2, \\ f_{xy}(x, y) &= -12, & f_{xx}(x, y) &= 6x, & f_{yy}(x, y) &= 48y. \end{aligned}$$

Los puntos extremos (x_0, y_0) de f deben satisfacer el sistema de ecuaciones

$$\begin{aligned} 3x_0^2 - 12y_0 &= 0, \\ -12x_0 + 24y_0^2 &= 0. \end{aligned}$$

La primera ecuación implica que

$$y_0 = \frac{1}{4}x_0^2.$$

Insertando esto en la segunda ecuación se tiene que

$$-12x_0 + \frac{24}{16}x_0^4 = 0 \iff x_0^4 - 8x_0 = 0.$$

Esta ecuación tiene solamente las soluciones $x_0^{(1)} = 0$ y $x_0^{(2)} = 2$, es decir los únicos puntos que hay que examinar son $(0, 0)$ y $(2, 1)$. Pero

$$\delta(0, 0) = f_{xx}(0, 0)f_{yy}(0, 0) - (f_{xy}(0, 0))^2 = -144 < 0,$$

por lo tanto f no posee un extremo en $(0, 0)$, mientras que

$$\delta(2, 1) = f_{xx}(2, 1)f_{yy}(2, 1) - (f_{xy}(2, 1))^2 = 432 > 0,$$

es decir, f posee un extremo local en $(2, 1)$; puesto que $f_{xx}(2, 1) = 12 > 0$, se trata de un mínimo local.

Ejemplo 3.12. Consideremos la función $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ dada por

$$f(x, y) = 3xy - x^2y - xy^2, \quad D(f) = \mathbb{R}^2.$$

Queremos determinar todos los extremos de f sobre el conjunto

$$M = \{(x, y) \mid x \geq 0, 0 \leq y \leq -x + 3\}.$$

Obtenemos aquí las siguientes condiciones necesarias para un extremo:

$$f_x = y(3 - y - 2x) = 0,$$

$$f_y = x(3 - x - 2y) = 0.$$

Las únicas soluciones de este sistema de ecuaciones son los puntos $(0, 0)$, $(0, 3)$, $(3, 0)$ y $(1, 1)$. Se tiene que $f(1, 1) = 1$. Los puntos $(0, 0)$, $(0, 3)$ y $(3, 0)$ están localizados en la frontera de M , donde $f(x, y) \equiv 0$. Obtenemos que f posee un máximo relativo en $(1, 1)$. Dado que M no puede contener ningún otro máximo de f , este máximo incluso es el máximo absoluto de f sobre M , es decir

$$0 \leq f(x, y) = 3xy - x^2y - xy^2 \leq 1 \quad \text{para todo } (x, y) \in M.$$

Este ejemplo muestra que a veces es posible determinar todos los extremos de f utilizando solamente la condición necesaria $\nabla f = \vec{0}$ y la forma analítica de f .

Ejemplo 3.13 (Problema 3, Certamen 1, Curso 2009/I). Sean f y g funciones $f, g : \mathbb{R} \rightarrow \mathbb{R}$ con $f \in C^2(a, b)$, $g \in C^2(c, d)$ para intervalos abiertos con $a < b$, $c < d$. Sean $f'' \neq 0$ sobre (a, b) y $g'' \neq 0$ sobre (c, d) . Supongamos que f tiene un extremo relativo (mínimo o máximo) en $x_0 \in (a, b)$, y que g tiene un extremo relativo del mismo tipo en $y_0 \in (c, d)$. ¿Se puede afirmar (demostración o contraejemplo) que también $\Phi : \mathbb{R}^2 \rightarrow \mathbb{R}$ posee en (x_0, y_0) un extremo del mismo tipo que f en x_0 y g en y_0 si (a) $\Phi(x, y) = f(x) + g(y)$, (b) $\Phi(x, y) = f(x) \cdot g(y)$?

Solución sugerida.

- a) Supongamos que f y g poseen un extremo relativo en $x_0 \in (a, b)$ e $y_0 \in (c, d)$, respectivamente. Entonces sabemos que $\Phi_x(x_0, y_0) = f'(x_0) = 0$ y $\Phi_y(x_0, y_0) = g'(y_0) = 0$, es decir (x_0, y_0) es un punto crítico, o sea $\nabla \Phi(x_0, y_0) = 0$. Luego obtenemos las segundas derivadas $\Phi_{xx} = f''(x)$, $\Phi_{yy} = g''(y)$ y $\Phi_{xy} = \Phi_{yx} = 0$, entonces

$$\delta(x_0, y_0) = \det \begin{bmatrix} \Phi_{xx}(x_0, y_0) & \Phi_{xy}(x_0, y_0) \\ \Phi_{yx}(x_0, y_0) & \Phi_{yy}(x_0, y_0) \end{bmatrix} = \begin{vmatrix} f''(x_0) & 0 \\ 0 & g''(y_0) \end{vmatrix} = f''(x_0)g''(y_0) > 0.$$

Si f y g poseen un mínimo relativo en $x_0 \in (a, b)$ e $y_0 \in (c, d)$, respectivamente, entonces $\Phi_{xx}(x_0, y_0) > 0$ y $\delta(x_0, y_0) > 0$, por lo tanto Φ posee un mínimo relativo en (x_0, y_0) . Asimismo, Si f y g poseen un máximo relativo en $x_0 \in (a, b)$ e $y_0 \in (c, d)$, respectivamente, entonces $\Phi_{xx}(x_0, y_0) < 0$ y $\delta(x_0, y_0) > 0$, por lo tanto Φ posee un máximo relativo en (x_0, y_0) .

- b) La afirmación es errónea. Contraejemplo: sean $a = c = -1$, $b = d = 1$, $f(x) = -x^2$, $g(y) = -y^2$. Las funciones f y g poseen un máximo relativo en $x_0 \in (a, b)$ e $y_0 \in (c, d)$, respectivamente, pero $\Phi(x, y) = x^2y^2$ tiene un mínimo en $(x_0, y_0) = (0, 0)$.

3.5. Extremos con restricciones y multiplicadores de Lagrange

Volveremos a la discusión de los extremos de una función de varias variables. El Teorema de las Funciones Implícitas nos permite el tratamiento de extremos de funciones sujetos a restricciones.

En muchas aplicaciones se presente el problema de que no solamente queremos estudiar una función $f : \mathbb{R}^n \rightarrow \mathbb{R}$, sino que se plantan una o varias restricciones adicionales a las cuales las variables están sujetas. Las restricciones están dadas en la forma de un sistema de ecuaciones

$$\begin{aligned} g_1(x_1, \dots, x_n) &= g_1(x) = 0, \\ &\vdots \\ g_m(x_1, \dots, x_n) &= g_m(x) = 0 \end{aligned}$$

o brevemente

$$g(x) = 0, \quad g = (g_1, \dots, g_m) : \mathbb{R}^n \rightarrow \mathbb{R}^m.$$

Se dice que la función f posee en x^0 un máximo local sujeto a las restricciones definidas por la función g si existe una vecindad $U(x^0)$ de x^0 tal que se tiene que $f(x) \leq f(x^0)$ para todo $x \in U(x^0)$ tal que $g_1(x) = 0, \dots, g_m(x) = 0$. Una definición análoga es válida para un mínimo sujeto a una restricción.

Para la determinación de los extremos de una función diferenciable $f : \mathbb{R}^n \rightarrow \mathbb{R}$ sin restricción nos interesan solamente aquellos puntos x^0 que satisfacen $\nabla f(x^0) = \vec{0}$. Hay una caracterización similar de aquellos puntos que son candidatos a ser extremo de f sujeto a la restricción $g(x) = 0$.

Teorema 3.12. *Se consideran las funciones $f : \mathbb{R}^n \rightarrow \mathbb{R}$ y $g = (g_1, \dots, g_m) : \mathbb{R}^n \rightarrow \mathbb{R}^m$ con $m < n$. Sea $x^0 = (x_1^0, \dots, x_n^0) \in D(f) \cap D(g)$, y sobre una vecindad $U(x^0)$ de x^0 sea $f \in C^1(U(x^0))$ y $g_\mu \in C^1(U(x^0))$ para $\mu = 1, \dots, m$. Sea el rango de la matriz*

$$\left(\frac{\partial g_\mu}{\partial x_\nu}(x^0) \right)_{\substack{\mu=1, \dots, m \\ \nu=1, \dots, n}}$$

igual m . Supongamos que la función f posee un extremo local en x^0 bajo la restricción $g(x) = 0$. Entonces existen constantes $\lambda_1, \dots, \lambda_m$, llamadas multiplicadores de Lagrange, tales que

$$\nabla f(x^0) = \lambda_1 \nabla g_1(x^0) + \dots + \lambda_m \nabla g_m(x^0).$$

Demostración.

1. Sin pérdida de la generalidad podemos suponer que

$$\begin{vmatrix} \frac{\partial g_1}{\partial x_1}(x^0) & \dots & \frac{\partial g_1}{\partial x_m}(x^0) \\ \vdots & & \vdots \\ \frac{\partial g_m}{\partial x_1}(x^0) & \dots & \frac{\partial g_m}{\partial x_m}(x^0) \end{vmatrix} \neq 0. \quad (3.12)$$

Esto significa que según el Teorema 3.2 podemos localmente despejar $x_1, \dots, x_m$ del sistema de ecuaciones

$$\begin{aligned} g_1(x_1, \dots, x_n) &= 0, \\ &\vdots \\ g_m(x_1, \dots, x_n) &= 0. \end{aligned} \quad (3.13)$$

Entonces, existen un $\delta > 0$ y m funciones

$$\begin{aligned} x_1 &= \varphi_1(x_{m+1}, \dots, x_n), \\ &\vdots \\ x_m &= \varphi_m(x_{m+1}, \dots, x_n) \end{aligned}$$

que pertenecen a C^1 para $|x_\nu - x_\nu^0| < \delta$ ($\nu = m+1, \dots, n$) y que son solución del sistema (3.13), es decir

$$\begin{aligned} g_1(\varphi_1(x_{m+1}, \dots, x_n), \dots, \varphi_m(x_{m+1}, \dots, x_n), x_{m+1}, \dots, x_n) &= 0, \\ &\vdots \\ g_m(\varphi_1(x_{m+1}, \dots, x_n), \dots, \varphi_m(x_{m+1}, \dots, x_n), x_{m+1}, \dots, x_n) &= 0. \end{aligned} \quad (3.14)$$

Insertando estas funciones en f , obtenemos una función de las variables $x_{m+1}, \dots, x_n$:

$$F(x_{m+1}, \dots, x_n) := f(\varphi_1(x_{m+1}, \dots, x_n), \dots, \varphi_m(x_{m+1}, \dots, x_n), x_{m+1}, \dots, x_n).$$

Según la hipótesis, esta función posee un extremo local en $\xi^0 = (x_{m+1}^0, \dots, x_n^0)$, por lo tanto se tiene que

$$\nabla F(\xi^0) = \vec{0}.$$

Utilizando la regla de la cadena obtenemos que la siguiente ecuación es válida en ξ^0 :

$$\frac{\partial F}{\partial x_\nu}(\xi^0) = \sum_{i=1}^m \frac{\partial f}{\partial x_i}(x^0) \frac{\partial \varphi_i}{\partial x_\nu}(\xi^0) + \frac{\partial f}{\partial x_\nu}(x^0) = 0, \quad \nu = m+1, \dots, n. \quad (3.15)$$

2. a) Debido a (3.12), el sistema

$$\sum_{\mu=1}^m \lambda_\mu \frac{\partial g_\mu}{\partial x_i}(x^0) = \frac{\partial f}{\partial x_i}(x^0), \quad i = 1, \dots, m \quad (3.16)$$

posee una solución única $(\lambda_1, \dots, \lambda_m)$.

b) De (3.14) obtenemos derivando la μ -ésima ecuación con respecto a x_ν :

$$\begin{aligned} \sum_{i=1}^m \frac{\partial g_\mu}{\partial x_i}(x^0) \frac{\partial \varphi_i}{\partial x_\nu}(\xi^0) + \frac{\partial g_\mu}{\partial x_\nu}(x^0) &= 0, \\ \nu = m+1, \dots, n, \quad \mu &= 1, \dots, m. \end{aligned} \quad (3.17)$$

De (3.15) y (3.16) se tiene que

$$\frac{\partial f}{\partial x_\nu}(x^0) = - \sum_{i=1}^m \frac{\partial f}{\partial x_i}(x^0) \frac{\partial \varphi_i}{\partial x_\nu}(\xi^0)$$

$$\begin{aligned}
&= - \sum_{i=1}^m \left(\sum_{\mu=1}^m \lambda_{\mu} \frac{\partial g_{\mu}}{\partial x_i}(x^0) \right) \frac{\partial \varphi_i}{\partial x_{\nu}}(\xi^0) \\
&= - \sum_{\mu=1}^m \lambda_{\mu} \left(\sum_{i=1}^m \frac{\partial g_{\mu}}{\partial x_i}(x^0) \frac{\partial \varphi_i}{\partial x_{\nu}}(\xi^0) \right).
\end{aligned}$$

En virtud de (3.17), esto implica que

$$\frac{\partial f}{\partial x_{\nu}}(x^0) = \sum_{\mu=1}^m \lambda_{\mu} \frac{\partial g_{\mu}}{\partial x_{\nu}}(x^0), \quad \nu = m+1, \dots, n.$$

Resumiendo (a) y (b) obtenemos

$$\frac{\partial f}{\partial x_{\nu}}(x^0) = \sum_{\mu=1}^m \lambda_{\mu} \frac{\partial g_{\mu}}{\partial x_{\nu}}(x^0), \quad \nu = 1, \dots, n,$$

lo que significa que

$$\nabla f(x^0) = \sum_{\mu=1}^m \lambda_{\mu} \nabla g_{\mu}(x^0).$$

■

Comentamos que este teorema entrega solamente un criterio necesario. Es decir, con su ayuda podemos solamente determinar aquellos puntos $(x_1^0, \dots, x_n^0)$ donde podría existir un extremo local de f sujeto a la restricción

$$g_1(x_1, \dots, x_n) = 0, \tag{3.18}$$

$$\vdots \tag{3.19}$$

$$g_m(x_1, \dots, x_n) = 0. \tag{3.20}$$

Para tal efecto se forman las $n + m$ ecuaciones

$$\begin{aligned}
&g_{\mu}(x_1^0, \dots, x_n^0) = 0, \quad \mu = 1, \dots, m, \\
&\frac{\partial f}{\partial x_{\nu}}(x_1^0, \dots, x_n^0) = \lambda_1 \frac{\partial g_1}{\partial x_{\nu}}(x_1^0, \dots, x_n^0) + \dots + \lambda_m \frac{\partial g_m}{\partial x_{\nu}}(x_1^0, \dots, x_n^0), \quad \nu = 1, \dots, n,
\end{aligned}$$

la cuales permiten determinar $\lambda_1, \dots, \lambda_m$ y $x_1^0, \dots, x_n^0$.

Sin embargo, para el tratamiento de extremos con restricciones no existen criterios simples suficientes (tales como para problemas de extremos sin restricciones) para poder decidir si en un punto $(x_1^0, \dots, x_n^0)$ efectivamente se tiene un extremo local de f sujeto a la restricción (3.18). Hay que decidir esto considerando la forma particular de f o utilizando consideraciones geométricas.

Ejemplo 3.14. *Queremos estudiar los extremos de la función*

$$f(x, y, z) = x + y + z$$

bajo las restricciones

$$g_1(x, y, z) = x^2 + y^2 - 2 = 0,$$

$$g_2(x, y, z) = x + z - 1 = 0.$$

Según el Teorema 3.12, los puntos críticos (candidatos a extremo) son aquellos puntos (x_0, y_0, z_0) que satisfacen

$$g_1(x_0, y_0, z_0) = 0,$$

$$g_2(x_0, y_0, z_0) = 0,$$

$$\nabla f(x_0, y_0, z_0) = \lambda_1 \nabla g_1(x_0, y_0, z_0) + \lambda_2 \nabla g_2(x_0, y_0, z_0).$$

Aquí obtenemos las cinco ecuaciones

$$x_0^2 + y_0^2 - 2 = 0, \tag{3.21}$$

$$x_0 + z_0 - 1 = 0, \tag{3.22}$$

$$1 = \lambda_1 \cdot 2x_0 + \lambda_2 \cdot 1, \tag{3.23}$$

$$1 = \lambda_1 \cdot 2y_0 + \lambda_2 \cdot 0, \tag{3.24}$$

$$1 = \lambda_1 \cdot 0 + \lambda_2 \cdot 1. \tag{3.25}$$

De estas ecuaciones debemos determinar x_0 , y_0 , z_0 , λ_1 y λ_2 . De (3.25) obtenemos $\lambda_2 = 1$, por lo tanto (3.23) entrega que $2\lambda_1 x_0 = 0$ y (3.24) implica que $2\lambda_1 y_0 = 1$, es decir $\lambda_1 \neq 0$ y por lo tanto $x_0 = 0$, luego $y_0 = \sqrt{2}$ o $y_0 = -\sqrt{2}$ y $z_0 = 1$. Los puntos críticos son $(0, \sqrt{2}, 1)$ y $(0, -\sqrt{2}, 1)$. En este caso, el punto $(0, \sqrt{2}, 1)$ corresponde a un máximo y el punto $(0, -\sqrt{2}, 1)$ a un mínimo.

Capítulo 4

Cálculo integral de funciones de varias variables I: teoría de la integración n -dimensional

4.1. Notación; sumas superiores e inferiores

Un concepto muy importante en el desarrollo del cálculo integral es la *partición* de un intervalo. Si $I = [a, b] \subset \mathbb{R}$ es un intervalo cerrado, entonces un conjunto de puntos

$$\{x_0, x_1, \dots, x_m\}$$

con $a = x_0 < x_1 < \dots < x_m = b$ se llama una *partición por puntos* del intervalo I . Para

$$I_k := [x_{k-1}, x_k], \quad k = 1, \dots, m$$

el conjunto

$$\{I_1, I_2, \dots, I_m\}$$

de estos intervalos cerrados define una *partición por intervalos* de $[a, b]$ con las propiedades

$$I = \bigcup_{k=1}^m I_k$$

y $I_k^0 \cap I_l^0 = \emptyset$ si $k \neq l$, donde I_k^0 denota el conjunto de los puntos interiores del intervalo I_k . Ver Figura 4.1.

Definición 4.1. Sean $-\infty < a_i < b_i < \infty$, $i = 1, \dots, n$, y sea

$$I = [a, b] = \{x = (x_1, \dots, x_n) \in \mathbb{R}^n \mid a_i \leq x_i \leq b_i, i = 1, \dots, n\}$$

un intervalo cerrado (ver Definición 1.3).

1. El número

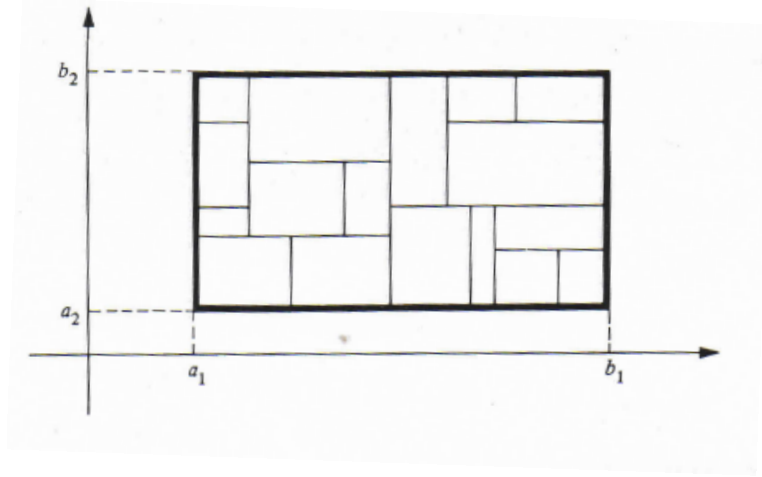
$$\mu(I) := \prod_{i=1}^n (b_i - a_i)$$

se llama medida n -dimensional o contenido n -dimensional de I .

2. Un conjunto $P = \{I_1, \dots, I_m\}$ de intervalos cerrados se llama *partición* de I si

$$I = \bigcup_{k=1}^m I_k$$

y $I_k^0 \cap I_l^0 = \emptyset$ si $k \neq l$, donde I_k^0 denota el conjunto de los puntos interiores del intervalo I_k .

FIGURA 4.1. Una partición del intervalo $[a_1, b_1] \times [a_2, b_2]$.

3. Sea $\delta(I_k)$ el diámetro de I_k , entonces la cantidad

$$\|P\| := \max_{1 \leq k \leq m} \delta(I_k)$$

se llama la norma de la partición P .

Teorema 4.1. Si $P = \{I_1, \dots, I_m\}$ es una partición de I , entonces

$$\mu(I) = \sum_{k=1}^m \mu(I_k).$$

Demostración. Tarea. ■

Definición 4.2. Se dice que una partición $P' = \{I'_1, \dots, I'_{m'}\}$ de I es un refinamiento de la partición $P = \{I_1, \dots, I_m\}$ de I si para cada $I'_{k'}$, $k' = 1, \dots, m'$ existe un I_k , $k = 1, \dots, m$, tal que $I'_{k'} \subset I_k$.

Comentamos que si P' es un refinamiento de P , entonces $\|P'\| \leq \|P\|$.

Definición 4.3. Sea la función $f : \mathbb{R}^n \rightarrow \mathbb{R}$ acotada sobre el intervalo cerrado $I \subset D(f)$, y sea $P = \{I_1, \dots, I_m\}$ una partición de I .

1. Definimos las cantidades

$$m_k(f) := \inf_{I_k} f(x), \quad M_k(f) := \sup_{I_k} f(x), \quad m(f) := \inf_I f(x), \quad M(f) := \sup_I f(x).$$

2. El siguiente número se llama suma inferior de f con respecto a P :

$$\underline{S}_P(f) := \sum_P m_k(f) \mu(I_k) = \sum_{k=1}^m m_k(f) \mu(I_k).$$

3. El siguiente número se llama suma superior de f con respecto a P :

$$\bar{S}_P(f) := \sum_P M_k(f) \mu(I_k) = \sum_{k=1}^m M_k(f) \mu(I_k).$$

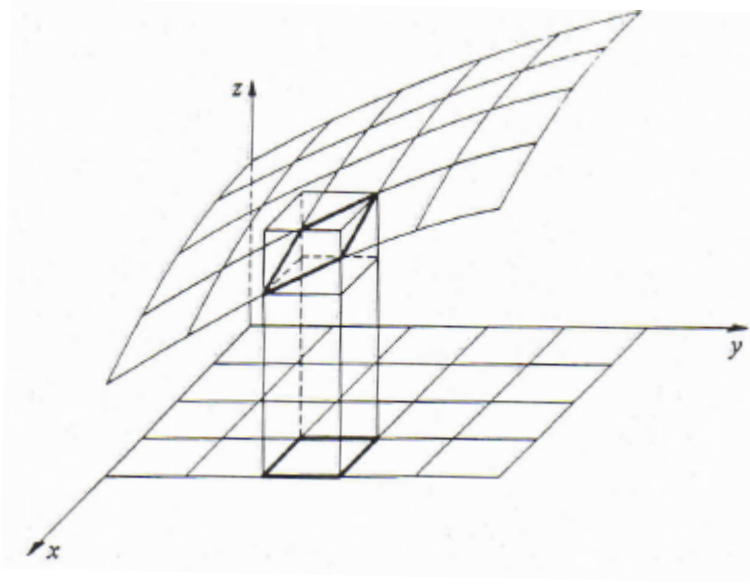


FIGURA 4.2. Ilustración de las sumas inferiores y superiores para una función f positiva, $n = 2$.

Teorema 4.2. *Para cada partición se tiene que*

$$m(f)\mu(I) \leq \underline{S}_P(f) \leq \bar{S}_P(f) \leq M(f)\mu(I).$$

Demostración. Puesto que $m(f) \leq m_k(f) \leq M_k(f) \leq M(f)$ para $k = 1, \dots, m$, obtenemos el enunciado después de multiplicar por $\mu(I_k)$ y sumando los productos. ■

Los siguientes resultados son análogos a los enunciados para el cálculo de funciones de una variable. Presentamos estos teoremas sin demostración.

Teorema 4.3. *Sea P' un refinamiento de P . Entonces*

1. $\bar{S}_{P'}(f) \leq \bar{S}_P(f)$,
2. $\underline{S}_{P'}(f) \geq \underline{S}_P(f)$.

Teorema 4.4. *Sean P_1 y P_2 particiones de I , entonces*

$$\underline{S}_{P_1}(f) \leq \bar{S}_{P_2}(f).$$

4.2. Integrales de Riemann-Darboux inferiores y superiores

Los Teoremas 4.2 y 4.3 implican la existencia de las integrales de Riemann-Darboux inferiores y superiores.

Definición 4.4. *Sea la función $f : \mathbb{R}^n \rightarrow \mathbb{R}$ acotada sobre el intervalo cerrado I .*

1. *La siguiente expresión se llama integral de Riemann-Darboux inferior de la función f sobre el intervalo I :*

$$\int_I f(x) dx = \int_I f(x_1, \dots, x_n) d(x_1, \dots, x_n) = \sup_P \underline{S}_P(f).$$

2. La siguiente expresión se llama integral de Riemann-Darboux superior de la función f sobre el intervalo I :

$$\int_I^{\bar{}} f(x) \, dx = \int_I^{\bar{}} f(x_1, \dots, x_n) \, d(x_1, \dots, x_n) = \inf_P \bar{S}_P(f).$$

Utilizando el Teorema 4.4 podemos deducir el siguiente resultado.

Teorema 4.5. *Se tiene que*

$$\int_I^{\bar{}} f(x) \, dx \leq \int_I f(x) \, dx.$$

Demostración. Tarea. ■

La Definición 4.3 tiene como consecuencia que para todo $\varepsilon > 0$ existe una partición P tal que

$$\underline{S}_P(f) > \int_I f(x) \, dx - \varepsilon; \quad \bar{S}_P(f) < \int_I^{\bar{}} f(x) \, dx + \varepsilon.$$

Sin embargo no es obvio que esta relación también es válida también para las particiones que poseen una norma suficientemente pequeña.

Teorema 4.6. *Para cada $\varepsilon > 0$ existe un $\delta_\varepsilon > 0$ tal que para todas las particiones P de I con $\|P\| < \delta_\varepsilon$ se tiene que*

$$\begin{aligned} \int_I f(x) \, dx - \varepsilon &< \underline{S}_P(f) \leq \int_I f(x) \, dx, \\ \int_I^{\bar{}} f(x) \, dx &\leq \bar{S}_P(f) < \int_I^{\bar{}} f(x) \, dx + \varepsilon. \end{aligned}$$

4.3. La integral de Riemann para intervalos

Mediante las integrales de Riemann-Darboux definiremos ahora la integral de Riemann para intervalos cerrados en $\mathbb{R}^n$.

Definición 4.5. *Sea la función $f : \mathbb{R}^n \rightarrow \mathbb{R}$ acotada sobre el intervalo cerrado $I \subset \mathbb{R}^n$. Si se tiene que*

$$\int_I f(x) \, dx = \int_I^{\bar{}} f(x) \, dx,$$

entonces f se llama Riemann-integrable sobre I . El valor común de las integrales superior e inferior se llama integral de Riemann de f sobre I , denotada

$$\int_I f(x) \, dx \quad \text{o} \quad \int_I f(x_1, \dots, x_n) \, d(x_1, \dots, x_n).$$

Para funciones $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ y $f : \mathbb{R}^3 \rightarrow \mathbb{R}$ también escribimos

$$\iint_I f(x, y) \, d(x, y) \quad \text{y} \quad \iiint_I f(x, y, z) \, d(x, y, z),$$

respectivamente.

Ejemplo 4.1. Sea $I \subset \mathbb{R}^n$ un intervalo cerrado. Sea la función $f : \mathbb{R}^n \rightarrow \mathbb{R}$ definida por

$$f(x) = f(x_1, \dots, x_n) := \begin{cases} 1 & \text{si } x_1 \in \mathbb{Q}, \\ 0 & \text{sino.} \end{cases}$$

Para cada partición P de I se tiene entonces que

$$\begin{aligned} \underline{S}_P(f) &= \sum_P \inf_{I_k} f(x) \mu(I_k) = 0, \\ \bar{S}_P(f) &= \sum_P \sup_{I_k} f(x) \mu(I_k) = \sum_P \mu(I_k) = \mu(I) > 0, \end{aligned}$$

por lo tanto f no puede ser Riemann-integrable.

Sin embargo, tal como para funciones de una variable se tiene que cada función continua sobre un intervalo cerrado es Riemann-integrable.

Teorema 4.7. Sea la función $f : \mathbb{R}^n \rightarrow \mathbb{R}$ continua sobre el intervalo cerrado $I \subset \mathbb{R}^n$. Entonces f es Riemann-integrable sobre I .

Demostración. Puesto que I es compacto, la función f es uniformemente continua sobre I , por lo tanto para $\varepsilon > 0$ existe un $\delta_\varepsilon > 0$ tal que para todo $x', x'' \in I$ con $d(x', x'') < \delta_\varepsilon$ se tiene que

$$|f(x') - f(x'')| < \frac{\varepsilon}{\mu(I)}.$$

Sea ahora $P = \{I_1, \dots, I_m\}$ una partición de I tal que $\|P\| < \delta_\varepsilon$. Entonces

$$\begin{aligned} 0 &\leq \int_I \bar{f}(x) \, dx - \int_I f(x) \, dx \\ &\leq \bar{S}_P(f) - \underline{S}_P(f) \\ &= \sum_{k=1}^m M_k(f) \mu(I_k) - \sum_{k=1}^m m_k(f) \mu(I_k) \\ &= \sum_{k=1}^m \left(\max_{I_k} f(x) - \min_{I_k} f(x) \right) \mu(I_k) \\ &< \frac{\varepsilon}{\mu(I)} \sum_{k=1}^m \mu(I_k) < \varepsilon. \end{aligned}$$

■

Teorema 4.8. Sea $I \subset \mathbb{R}^n$ un intervalo cerrado y sea c una constante. Entonces

$$\int_I c \, dx = c \cdot \mu(I).$$

En particular, para $c = 1$,

$$\mu(I) = \int_I dx.$$

Demostración. Sea $P = \{I_1, \dots, I_m\}$ una partición arbitraria de I . Entonces

$$\bar{S}_P(f) = \sum_{k=1}^m c \cdot \mu(I_k) = c \sum_{k=1}^m \mu(I_k) = c \cdot \mu(I),$$

por lo tanto según el Teorema 4.7

$$\int_I c \, dx = c \cdot \mu(I).$$

■

4.4. Sumas de Riemann

Definición 4.6. Sea la función $f : \mathbb{R}^n \rightarrow \mathbb{R}$ acotada sobre el intervalo cerrado $I \subset \mathbb{R}^n$, y sea $P = \{I_1, \dots, I_m\}$ una partición de I ; además, sea $\xi = \{\xi_1, \dots, \xi_m\}$ un conjunto de puntos tales que $\xi_k \in I_k$ para $k = 1, \dots, m$. Entonces

$$S_P(f, \xi) = \sum_P f(\xi_k) \mu(I_k) = \sum_{k=1}^m f(\xi_k) \mu(I_k)$$

se llama suma de Riemann de f respecto a P .

Comentamos que las sumas superiores e inferiores, en general, no son sumas de Riemann porque no necesariamente los valores $M_k(f)$ y $m_k(f)$ deben ser asumidos por f sobre I_k . Además, tal como en la teoría de funciones de una variable se tiene ahora que

$$\underline{S}_P(f) \leq S_P(f, \xi) \leq \bar{S}_P(f).$$

La convergencia de las sumas de Riemann se define de manera análoga al cálculo de las funciones de una variable.

Definición 4.7. Sea la función f acotada sobre el intervalo cerrado I . Si existe un número $J \in \mathbb{R}$ y para cada $\varepsilon > 0$ existe un $\delta_\varepsilon > 0$ tal que para cada partición $P = \{I_1, \dots, I_m\}$ de I con $\|P\| < \delta_\varepsilon$, con $\xi_k \in I_k$ elegido arbitrariamente, se tiene que

$$|S_P(f, \xi) - J| < \varepsilon,$$

entonces se dice que las sumas de Riemann convergen a J , y escribimos

$$J = \lim_{\|P\| \rightarrow 0} S_P(f, \xi).$$

El próximo teorema, cuya demostración es análoga al cálculo de funciones de una variable y por lo tanto omitida, muestra que la convergencia de las sumas de Riemann es equivalente con la existencia de la integral de Riemann.

Teorema 4.9. *Sea $I \subset \mathbb{R}^n$ un intervalo cerrado.*

1. *Si f es Riemann-integrable sobre I , entonces existe el límite $\lim_{\|P\| \rightarrow 0} S_P(f, \xi)$, y se tiene que*

$$\lim_{\|P\| \rightarrow 0} S_P(f, \xi) = \int_I f(x) \, dx.$$

2. *Por otro lado, si existe el límite $\lim_{\|P\| \rightarrow 0} S_P(f, \xi)$, entonces f es Riemann-integrable sobre I y se tiene que*

$$\int_I f(x) \, dx = \lim_{\|P\| \rightarrow 0} S_P(f, \xi).$$

4.5. Integrales iteradas

Ahora introduciremos un método práctico para la computación de una integral n -dimensional. Basicamente reemplazaremos la integración n -dimensional por n integraciones unidimensionales.

Teorema 4.10. *Sea la función $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ Riemann-integrable sobre el rectángulo*

$$I := \{(x, y) \mid a \leq x \leq b, c \leq y \leq d\}.$$

1. *Supongamos que para cada $x \in [a, b]$ existe la integral*

$$\int_c^d f(x, y) \, dy.$$

Entonces existe la integral iterada

$$\int_a^b \left[\int_c^d f(x, y) \, dy \right] dx,$$

y se tiene que

$$\iint_I f(x, y) \, d(x, y) = \int_a^b \left[\int_c^d f(x, y) \, dy \right] dx.$$

2. *Supongamos que para cada $y \in [c, d]$ existe la integral*

$$\int_a^b f(x, y) \, dx.$$

Entonces existe la integral iterada

$$\int_c^d \left[\int_a^b f(x, y) \, dx \right] dy,$$

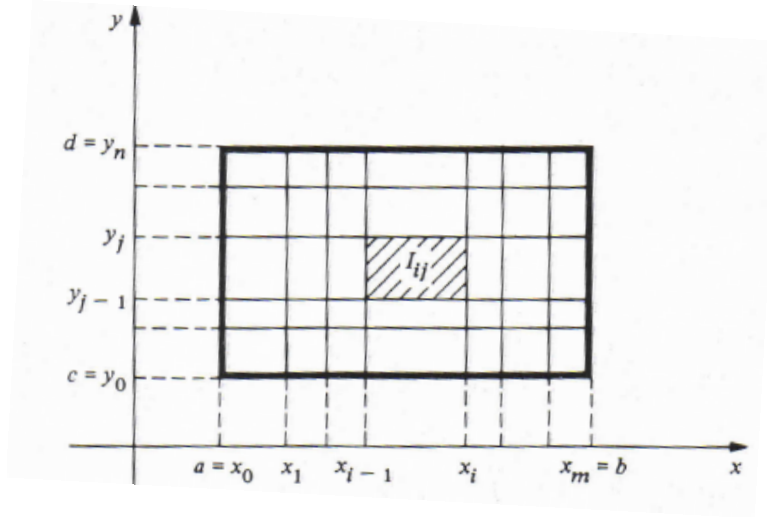


FIGURA 4.3. Ilustración de la demostración del Teorema 4.10.

y se tiene que

$$\iint_I f(x, y) \, d(x, y) = \int_c^d \left[\int_a^b f(x, y) \, dx \right] dy.$$

Demostración. Consideremos las siguientes particiones por puntos de los intervalos respectivos $[a, b]$ y $[c, d]$:

$$P_x = \{x_0, \dots, x_m\}, \quad P_y = \{y_0, \dots, y_n\}.$$

Así, el conjunto de los rectángulos

$$I_{ij} := \{(x, y) \mid x_{i-1} \leq x \leq x_i, y_{j-1} \leq y \leq y_j\}$$

forma una partición

$$P = \{I_{ij}\}_{\substack{i=1, \dots, m \\ j=1, \dots, n}}$$

del rectángulo I (ver Figura 4.3). Como siempre, definimos

$$m_{ij} := \inf_{I_{ij}} f(x, y), \quad M_{ij} := \sup_{I_{ij}} f(x, y).$$

Ahora, según el Teorema 4.6, podemos elegir para un $\varepsilon > 0$ dado un $\delta_\varepsilon > 0$ tal que para todas particiones P_x y P_y tales que $\|P_x\| < \delta_\varepsilon$ y $\|P_y\| < \delta_\varepsilon$ se tiene que

$$\underline{S}_P(f) > \iint_I f(x, y) \, d(x, y) - \varepsilon, \quad \bar{S}_P(f) < \iint_I f(x, y) \, d(x, y) + \varepsilon. \quad (4.1)$$

1. Consideremos sobre el intervalo $[a, b]$ la función $F : \mathbb{R} \rightarrow \mathbb{R}$ definida por

$$F(x) := \int_c^d f(x, y) \, dy.$$

Esta función es acotada sobre $[a, b]$, y podemos formar la suma de Riemann (en el sentido del cálculo de funciones de una variable)

$$S_{P_x}(F, \xi) = \sum_{i=1}^m F(\xi_i)(x_i - x_{i-1}), \quad \text{donde } \xi_i \in [x_{i-1}, x_i].$$

Evidentemente se tiene que

$$m_{ij} \leq f(\xi_i, y) \leq M_{ij} \quad \text{para todo } y \in [y_{j-1}, y_j].$$

Ahora, integrando sobre $[y_{j-1}, y_j]$ obtenemos

$$m_{ij}(y_j - y_{j-1}) \leq \int_{y_{j-1}}^{y_j} f(\xi_i, y) dy \leq M_{ij}(y_j - y_{j-1}).$$

Multiplicando por $(x_i - x_{i-1})$, sumando sobre j y luego sobre i obtenemos

$$\begin{aligned} \underline{S}_P(f) &= \sum_{i=1}^m \sum_{j=1}^n m_{ij} \mu(I_{ij}) \\ &\leq \sum_{i=1}^m \left[\int_c^d f(\xi_i, y) dy \right] (x_i - x_{i-1}) \\ &= \sum_{i=1}^m F(\xi_i)(x_i - x_{i-1}) \\ &\leq \sum_{i=1}^m \sum_{j=1}^n M_{ij} \mu(I_{ij}) = \bar{S}_P(f). \end{aligned}$$

Concluimos que en virtud de (4.1) para cada P_x tal que $\|P_x\| < \delta_\varepsilon$ y cada selección de los ξ_i

$$\iint_I f(x, y) d(x, y) - \varepsilon \leq S_{P_x}(F, \xi) \leq \iint_I f(x, y) d(x, y) + \varepsilon,$$

lo que nos permite concluir que la afirmación (1) del teorema es válida.

2. La demostración de (2) es análoga. ■

Comentamos que la existencia de las integrales iteradas

$$\int_a^b \left[\int_c^d f(x, y) dy \right] dx \quad \text{y} \quad \int_c^d \left[\int_a^b f(x, y) dx \right] dy \quad (4.2)$$

no necesariamente implica la existencia de

$$\iint_I f(x, y) d(x, y). \quad (4.3)$$

Vice versa, la existencia de (4.3) no necesariamente implica la existencia de las integrales iteradas (4.2).

Si la función f es continua sobre

$$I := \{(x, y) \mid a \leq x \leq b, c \leq y \leq d\},$$

entonces para cada $x \in [a, b]$ y cada $y \in [c, d]$ existen las respectivas integrales

$$\int_c^d f(x, y) dy \quad \text{y} \quad \int_a^b f(x, y) dx,$$

además f es Riemann-integrable (según el Teorema 4.7) sobre I . Entonces, en virtud del Teorema 4.10 se tiene que

$$\int_c^d \left[\int_a^b f(x, y) dx \right] dy = \iint_I f(x, y) d(x, y) = \int_a^b \left[\int_c^d f(x, y) dy \right] dx,$$

es decir podemos intercambiar el orden de las integraciones de las integrales iteradas (Teorema de Fubini).

Finalmente, comentamos que si la función f tiene la forma

$$f(x, y) = f_1(x)f_2(y),$$

donde f_1 es continua sobre $[a, b]$ y f_2 es continua sobre $[c, d]$ se tiene que

$$\iint_I f(x, y) d(x, y) = \int_a^b f_1(x) dx \cdot \int_c^d f_2(y) dy.$$

Ejemplo 4.2. Sea $I := \{(x, y) \mid 0 \leq x \leq 1, 1 \leq y \leq 2\}$, y sea f definida sobre I por $f(x, y) = x^y$. Puesto que f es continua sobre I , el Teorema 4.10 nos permite concluir que

$$\iint_I x^y d(x, y) = \int_1^2 \left[\int_0^1 x^y dx \right] dy.$$

Aquí obtenemos

$$\int_0^1 x^y dx = \left[\frac{x^{y+1}}{y+1} \right]_{x=0}^{x=1} = \frac{1}{y+1}$$

y por lo tanto

$$\iint_I x^y d(x, y) = \int_1^2 \frac{1}{1+y} dy = \ln \frac{3}{2}.$$

El resultado del Teorema 4.10 es válido para integrales n -dimensionales generales.

Teorema 4.11. Sea la función $f : \mathbb{R}^n \rightarrow \mathbb{R}$ Riemann-integrable sobre el intervalo

$$I := \{(x_1, \dots, x_n) \mid a_i \leq x_i \leq b_i, i = 1, \dots, n\},$$

además definimos para cada ν , $1 \leq \nu \leq n$

$$I_{x_\nu} := \{(x_1, \dots, x_{\nu-1}, x_{\nu+1}, \dots, x_n) \mid a_i \leq x_i \leq b_i, i = 1, \dots, n, i \neq \nu\}.$$

Entonces se tiene lo siguiente.

1. Si para cada $x_\nu \in [a_\nu, b_\nu]$ existe la integral

$$\int_{I_{x_\nu}} f(x_1, \dots, x_\nu, \dots, x_n) d(x_1, \dots, x_{\nu-1}, x_{\nu+1}, \dots, x_n),$$

también existe la integral iterada

$$\int_{a_\nu}^{b_\nu} \left[\int_{I_{x_\nu}} f(x_1, \dots, x_\nu, \dots, x_n) d(x_1, \dots, x_{\nu-1}, x_{\nu+1}, \dots, x_n) \right] dx_\nu,$$

y esta integral posee el valor

$$\int_I f(x) dx.$$

2. Si para cada $(x_1, \dots, x_{\nu-1}, x_{\nu+1}, \dots, x_n) \in I_{x_\nu}$ existe la integral

$$\int_{a_\nu}^{b_\nu} f(x_1, \dots, x_\nu, \dots, x_n) dx_\nu,$$

también existe la integral iterada

$$\int_{I_{x_\nu}} \left[\int_{a_\nu}^{b_\nu} f(x_1, \dots, x_\nu, \dots, x_n) dx_\nu \right] d(x_1, \dots, x_{\nu-1}, x_{\nu+1}, \dots, x_n)$$

y esta integral posee el valor

$$\int_I f(x) dx.$$

Demostración. La demostración es análoga a la del Teorema 4.10. ■

Ejemplo 4.3. Sea

$$I := \{(x, y, z) \mid 0 \leq x \leq 2, 0 \leq y \leq 1, 2 \leq z \leq 4\},$$

y sea f definida sobre I por

$$f(x, y, z) = x + y + z.$$

Puesto que f es continua sobre I , se tiene según el Teorema 4.11 para

$$I_z := \{(x, y) \mid 0 \leq x \leq 2, 0 \leq y \leq 1\}$$

que

$$\iiint_I (x + y + z) d(x, y, z) = \int_2^4 \left[\iint_{I_z} (x + y + z) d(x, y) \right] dz.$$

La integral interior puede ser descompuesta de la siguiente manera:

$$\iint_{I_z} (x + y + z) \, d(x, y) = \int_0^1 \left[\int_0^2 (x + y + z) \, dx \right] dy,$$

así que

$$\begin{aligned} \iiint_I (x + y + z) \, d(x, y, z) &= \int_0^4 \left[\int_0^1 \left(\int_0^2 (x + y + z) \, dx \right) dy \right] dz \\ &= \int_2^4 \left[\int_0^1 (2 + 2y + 2z) \, dy \right] dz = \int_2^4 (3 + 2z) \, dz = 18. \end{aligned}$$

4.6. Integrales de Riemann-Darboux inferiores y superiores sobre conjuntos acotados

Hasta ahora hemos solamente considerado el caso de una función f definida sobre un intervalo I . Para avanzar en el cálculo de varias variables es necesario desarrollar una teoría de integración que permite la integración de una función sobre subconjuntos de $\mathbb{R}^n$ más generales.

Definición 4.8. Sea $X \subset \mathbb{R}^n$ un conjunto acotado y sea la función $f : \mathbb{R}^n \rightarrow \mathbb{R}$ acotada sobre X . Entonces la función $f_X : \mathbb{R}^n \rightarrow \mathbb{R}$ con $D(f_X) = \mathbb{R}^n$ definida por

$$f_X(x) = \begin{cases} f(x) & \text{si } x \in X, \\ 0 & \text{si } x \notin X \end{cases}$$

se llama extensión de la función f a $\mathbb{R}^n$.

Definición 4.9. Sea $X \subset \mathbb{R}^n$ un conjunto acotado. Entonces la función $c_X : \mathbb{R}^n \rightarrow \mathbb{R}$ con $D(c_X) = \mathbb{R}^n$ definida por

$$c_X(x) = \begin{cases} 1 & \text{si } x \in X, \\ 0 & \text{si } x \notin X \end{cases}$$

se llama función característica de X .

Teorema 4.12. Sean I y K intervalos cerrados de $\mathbb{R}^n$ con $I \subset K$. Sea la función $f : \mathbb{R}^n \rightarrow \mathbb{R}$ acotada sobre I . En este caso se tiene que

$$\int_I f(x) \, dx = \int_K f_I(x) \, dx, \tag{4.4}$$

$$\int_I f(x) \, dx = \int_K f_I(x) \, dx. \tag{4.5}$$

Demostración. Demostraremos solamente (4.4); la demostración de (4.5) es análoga. Sea $\varepsilon > 0$ dado. Según el Teorema 4.6 podemos elegir $\delta_\varepsilon > 0'$ de tal manera que para todas particiones P_I de I y P_K de K con $\|P_I\| < \delta_\varepsilon$ y $\|P_K\| < \delta_\varepsilon$ se tiene que

$$\begin{aligned} \int_{\underline{I}} f(x) dx - \frac{\varepsilon}{3} &< \underline{S}_{P_I}(f) \leq \int_{\underline{I}} f(x) dx, \\ \int_{\underline{K}} f_I(x) dx - \frac{\varepsilon}{3} &\leq \underline{S}_{P_K}(f_I) \leq \int_{\underline{K}} f_I(x) dx. \end{aligned}$$

Sea ahora P'_K una partición de K con $\|P'_K\| < \delta_\varepsilon$ que contiene un subconjunto P'_I que a su vez es una partición de I . La suma inferior $\underline{S}_{P'_K}(f_I)$ está compuesta por la suma inferior $\underline{S}_{P'_I}(f)$ y ciertas contribuciones que provienen de los puntos de la frontera de I . Pero si elegimos P'_K suficientemente fina fuera de I podemos lograr que

$$|\underline{S}_{P'_K}(f_I) - \underline{S}_{P'_I}(f)| < \frac{\varepsilon}{3}.$$

Así se tiene que

$$\begin{aligned} &\left| \int_{\underline{I}} f(x) dx - \int_{\underline{K}} f_I(x) dx \right| \\ &\leq \int_{\underline{I}} f(x) dx - \underline{S}_{P'_I}(f) + \int_{\underline{K}} f_I(x) dx - \underline{S}_{P'_K}(f_I) + |\underline{S}_{P'_K}(f_I) - \underline{S}_{P'_I}(f)| \\ &\leq \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon. \end{aligned}$$

Esto concluye la demostración de (4.4). ■

Definición 4.10. Sea $X \subset \mathbb{R}^n$ un conjunto acotado, y sea la función $f : \mathbb{R}^n \rightarrow \mathbb{R}$ acotada sobre X ; además, sea $I \subset \mathbb{R}^n$ un intervalo acotado de $\mathbb{R}^n$ con $X \subset I$.

1. La siguiente expresión se llama integral de Riemann-Darboux inferior de f sobre X :

$$\int_{\underline{X}} f(x) dx = \int_{\underline{I}} f_X(x) dx.$$

2. La siguiente expresión se llama integral de Riemann-Darboux superior de f sobre X :

$$\int_{\overline{X}} f(x) dx = \int_{\overline{I}} f_X(x) dx.$$

Enfatizamos que las expresiones

$$\int_{\underline{X}} f(x) dx \quad \text{y} \quad \int_{\overline{X}} f(x) dx$$

son independientes del particular intervalo cerrado considerado $I \subset \mathbb{R}^n$ con $X \subset I$. Para ver eso, supongamos que $J \subset \mathbb{R}^n$ sea otro intervalo cerrado $X \subset J$. En este caso elegimos un intervalo cerrado $K \subset \mathbb{R}^n$ tal que $I \cup J \subset K$, y obtenemos las siguientes igualdades como consecuencia del Teorema 4.12:

$$\begin{aligned} \int_{\bar{X}} f(x) dx &= \int_{\bar{I}} f_X(x) dx = \int_{\bar{K}} f_X(x) dx = \int_{\bar{J}} f_X(x) dx, \\ \int_X f(x) dx &= \int_I f_X(x) dx = \int_K f_X(x) dx = \int_J f_X(x) dx. \end{aligned}$$

Queremos derivar algunas propiedades importantes de las integrales superiores e inferiores de Riemann-Darboux.

Teorema 4.13. *Sea $X \subset \mathbb{R}^n$ un conjunto acotado y sea $f : \mathbb{R}^n \rightarrow \mathbb{R}$ acotada sobre X , entonces*

$$\int_{\bar{X}} f(x) dx \leq \int_X f(x) dx.$$

Demostración. Si I es un intervalo acotado con $X \subset I$, entonces se tiene lo siguiente en virtud de la Definición 4.10 y del Teorema 4.5:

$$\int_{\bar{X}} f(x) dx = \int_{\bar{I}} f_X(x) dx \leq \int_I f_X(x) dx = \int_X f(x) dx.$$

■

Teorema 4.14. *Sea $X \subset \mathbb{R}^n$ un conjunto acotado y sean $f, g : \mathbb{R}^n \rightarrow \mathbb{R}$ acotadas sobre X , entonces*

$$\int_{\bar{X}} f(x) dx + \int_{\bar{X}} g(x) dx \leq \int_{\bar{X}} (f(x) + g(x)) dx \leq \int_X (f(x) + g(x)) dx \leq \int_X f(x) dx + \int_X g(x) dx.$$

Demostración. Sea I un intervalo cerrado tal que $X \subset I$ y sea $P = \{I_1, \dots, I_m\}$ una partición de I . Sean f_X y g_X las extensiones de f y g . Evidentemente $(f+g)_X = f_X + g_X$. Multiplicando

$$M_k(f_X + g_X) \leq M_k(f_X) + M_k(g_X)$$

por $\mu(I_k)$ y sumando sobre k obtenemos

$$\int_I (f_X(x) + g_X(x)) dx \leq \bar{S}_P(f_X + g_X) \leq \bar{S}_P(f_X) + \bar{S}_P(g_X).$$

Ahora para $\varepsilon > 0$ existe un $\delta_\varepsilon > 0$ tal que

$$0 \leq \bar{S}_P(f_X) - \int_I f_X(x) dx < \varepsilon, \quad 0 \leq \bar{S}_P(g_X) - \int_I g_X(x) dx < \varepsilon$$

para toda partición P tal que $\|P\| < \delta_\varepsilon$, luego

$$\int_I (f_X(x) + g_X(x)) \, dx < \int_I f_X(x) \, dx + \int_I g_X(x) \, dx + 2\varepsilon,$$

lo que demuestra la desigualdad derecha. La demostración de la desigualdad izquierda es análoga. ■

Teorema 4.15. *Sea $X \subset \mathbb{R}^n$ un conjunto acotado y sea $f : \mathbb{R}^n \rightarrow \mathbb{R}$ acotada sobre X . Entonces se tiene para cada $c \geq 0$*

$$\begin{aligned} \int_{\bar{X}} (c \cdot f(x)) \, dx &= c \cdot \int_{\bar{X}} f(x) \, dx, \\ \int_X (c \cdot f(x)) \, dx &= c \cdot \int_X f(x) \, dx. \end{aligned}$$

Demostración. Demostraremos solamente la primera identidad; la demostración de la segunda es análoga. Evidentemente, $(c \cdot f)_X = c \cdot f_X$ para todo $c \geq 0$. Sea I un intervalo cerrado tal que $X \subset I$. Entonces para cada partición P de I se tiene que

$$\underline{S}_P(c \cdot f_X) = \sum_P \left(\inf_{I_k} (c \cdot f_X(x)) \right) \mu(I_k) = c \cdot \sum_P \left(\inf_{I_k} f_X(x) \right) \mu(I_k) = c \cdot \underline{S}_P(f_X),$$

por lo tanto

$$\int_{\bar{X}} (c \cdot f(x)) \, dx = \sup_P \underline{S}_P(c \cdot f_X) = c \cdot \sup_P \underline{S}_P(f_X) = c \cdot \int_{\bar{X}} f(x) \, dx. \quad \blacksquare$$

Para la multiplicación de una función por (-1) se tiene el siguiente teorema.

Teorema 4.16. *Sea $X \subset \mathbb{R}^n$ un conjunto acotado y sea $f : \mathbb{R}^n \rightarrow \mathbb{R}$ acotada sobre X . Entonces se tiene que*

$$\begin{aligned} \int_{\bar{X}} (-f(x)) \, dx &= - \int_{\bar{X}} f(x) \, dx, \\ \int_X (-f(x)) \, dx &= - \int_X f(x) \, dx. \end{aligned}$$

Demostración. Demostraremos solamente la primera identidad; la demostración de la segunda es análoga. Evidentemente, $(-f)_X = -f_X$. Sea I un intervalo cerrado tal que $X \subset I$. Entonces para cada partición P de I se tiene que

$$\underline{S}_P(-f_X) = \sum_P \left(\inf_{I_k} (-f_X(x)) \right) \mu(I_k) = - \sum_P \left(\sup_{I_k} f_X(x) \right) \mu(I_k) = -\bar{S}_P(f_X),$$

por lo tanto

$$\int_{\bar{X}} (-f(x)) \, dx = \sup_P \underline{S}_P(-f_X) = -\inf_P \bar{S}_P(f_X) = -\int_X f(x) \, dx.$$

■

Teorema 4.17. *Sea $X \subset \mathbb{R}^n$ un conjunto acotado y sean $f, g : \mathbb{R}^n \rightarrow \mathbb{R}$ acotadas, además sea $f(x) \leq g(x)$ para $x \in X$. Entonces se tiene que*

$$\begin{aligned} \int_{\bar{X}} f(x) \, dx &\leq \int_{\bar{X}} g(x) \, dx, \\ \int_X f(x) \, dx &\leq \int_X g(x) \, dx. \end{aligned}$$

Demostración. Demostraremos solamente la primera desigualdad; la demostración de la segunda es análoga. Sea I un intervalo cerrado tal que $X \subset I$. Entonces para cada partición P de I se tiene que

$$\underline{S}_P(f_X) = \sum_P \left(\inf_{I_k} (f_X(x)) \right) \mu(I_k) \leq \sum_P \left(\inf_{I_k} (g_X(x)) \right) \mu(I_k) = \underline{S}_P(g_X),$$

por lo tanto

$$\int_{\bar{X}} f(x) \, dx = \sup_P \underline{S}_P(f_X) \leq \sup_P \underline{S}_P(g_X) = \int_{\bar{X}} g(x) \, dx.$$

■

Teorema 4.18. *Sean $X_1, X_2 \subset \mathbb{R}^n$ conjuntos acotados con $X_1 \cap X_2 = \emptyset$ y sea $f : \mathbb{R}^n \rightarrow \mathbb{R}$ acotada sobre $X_1 \cup X_2$, entonces se tiene que*

$$\begin{aligned} \int_{\bar{X}_1} f(x) \, dx + \int_{\bar{X}_2} f(x) \, dx &\leq \int_{X_1 \cup X_2} f(x) \, dx \leq \int_{\bar{X}_1} f(x) \, dx + \int_{\bar{X}_2} f(x) \, dx \\ &\leq \int_{X_1 \cup X_2} f(x) \, dx \leq \int_{\bar{X}_1} f(x) \, dx + \int_{\bar{X}_2} f(x) \, dx. \end{aligned}$$

Demostración. Demostraremos solamente la primera y la segunda desigualdad; las demostraciones de la tercera y de la cuarta son análogas. Consideremos las funciones

$$f_1(x) := \begin{cases} f(x) & \text{si } x \in X_1, \\ 0 & \text{si } x \notin X_1, \end{cases} \quad f_2(x) := \begin{cases} f(x) & \text{si } x \in X_2, \\ 0 & \text{si } x \notin X_2. \end{cases}$$

Estas funciones satisfacen $f = f_1 + f_2$ sobre $X_1 \cup X_2$; además, sea I un intervalo cerrado tal que $X_1 \cup X_2 \subset I$. Utilizando el Teorema 4.14 y la Definición 4.10 podemos concluir que

$$\int_{\bar{X}_1} f(x) \, dx + \int_{\bar{X}_2} f(x) \, dx = \int_I f_1(x) \, dx + \int_I f_2(x) \, dx \leq \int_I (f_1(x) + f_2(x)) \, dx$$

$$= \int_{\underline{X_1 \cup X_2}} (f_1(x) + f_2(x)) \, dx = \int_{\underline{X_1 \cup X_2}} f(x) \, dx,$$

lo que establece la primera desigualdad. Para demostrar la segunda desigualdad, notamos que los Teoremas 4.14 y 4.16 implican que

$$\begin{aligned} \int_{\underline{X_1 \cup X_2}} f(x) \, dx &= \int_{\underline{I}} (f_1(x) + f_2(x)) \, dx + \int_{\underline{I}} (-f_2(x)) \, dx - \int_{\underline{I}} (-f_2(x)) \, dx \\ &\leq \int_{\underline{I}} (f_1(x) + f_2(x) - f_2(x)) \, dx + \int_{\underline{I}} f_2(x) \, dx \\ &= \int_{\underline{X_1}} f(x) \, dx + \int_{\underline{X_2}} f(x) \, dx, \end{aligned}$$

lo que concluye la demostración de la segunda desigualdad. ■

4.7. La integral de Riemann sobre conjuntos acotados

Con la ayuda de las integrales de Riemann-Darboux superiores e inferiores podemos ahora definir la integral de Riemann para conjuntos acotados generales de $\mathbb{R}^n$.

Definición 4.11. Sea $X \subset \mathbb{R}^n$ un conjunto acotado y sea $f : \mathbb{R}^n \rightarrow \mathbb{R}$ acotada sobre X . Si se tiene que

$$\int_{\underline{X}} f(x) \, dx = \int_{\overline{X}} f(x) \, dx,$$

entonces f se llama Riemann-integrable sobre X . El valor común de ambas expresiones se llama la integral de Riemann de f sobre X y es denotado por

$$\int_X f(x) \, dx.$$

Si $X = I$ es un intervalo cerrado, entonces esta definición coincide con la Definición 4.5.

Si f es Riemann-integrable sobre X , entonces el Teorema 4.12 implica que para cada intervalo cerrado $I \supset X$ se tiene que

$$\int_X f(x) \, dx = \int_I f_X(x) \, dx.$$

Teorema 4.19 (Criterio de integrabilidad de Riemann). Sea $f : \mathbb{R}^n \rightarrow \mathbb{R}$ una función acotada. La función f es Riemann-integrable sobre el conjunto acotado X si y sólo si para cada $\varepsilon > 0$ y cada intervalo cerrado $I \subset \mathbb{R}^n$ con $X \subset I$ existe una partición P de I tal que

$$\bar{S}_P(f_X) - \underline{S}_P(f_X) < \varepsilon.$$

Demostración.

1. Sea I un intervalo cerrado con $X \subset I$, y supongamos que para $\varepsilon > 0$ existe una partición P de I tal que

$$\bar{S}_P(f_X) - \underline{S}_P(f_X) < \varepsilon;$$

en este caso se tiene que

$$0 \leq \int_I f_X(x) dx - \int_{\bar{I}} f_X(x) dx \leq \bar{S}_P(f_X) - \underline{S}_P(f_X) < \varepsilon.$$

Puesto que $\varepsilon > 0$ es arbitrario, se tiene que

$$\int_I f_X(x) dx = \int_{\bar{I}} f_X(x) dx,$$

por lo tanto, según la Definición 4.10 también se tiene que

$$\int_X f(x) dx = \int_{\bar{X}} f(x) dx,$$

y concluimos que según la Definición 4.11, la función f es Riemann-integrable sobre X .

2. Ahora sea f Riemann-integrable sobre X , y sea I un intervalo cerrado con $X \subset I$, y sea $\varepsilon > 0$ dado. Entonces existen particiones P_1 y P_2 de I tales que

$$\bar{S}_{P_1}(f_X) < \int_I f_X(x) dx + \frac{\varepsilon}{2}, \quad \underline{S}_{P_2}(f_X) > \int_I f_X(x) dx - \frac{\varepsilon}{2}.$$

Sea ahora P una partición de I que sea un refinamiento tanto de P_1 como de P_2 . Entonces según el Teorema 4.3 se tiene que

$$\bar{S}_P(f_X) - \underline{S}_P(f_X) \leq \bar{S}_{P_1}(f_X) - \underline{S}_{P_2}(f_X) < \varepsilon.$$

■

En analogía con la teoría de funciones Riemann-integrables de una variable tenemos los siguientes teoremas.

Teorema 4.20. *Sea la función $f : \mathbb{R}^n \rightarrow \mathbb{R}$ Riemann-integrable sobre el conjunto X , y sea $c \in \mathbb{R}$. Entonces también la función $c \cdot f$ es Riemann-integrable sobre X , y se tiene que*

$$\int_X (c \cdot f(x)) dx = c \cdot \int_X f(x) dx.$$

Demostración. Para $c \geq 0$, el enunciado es una consecuencia del Teorema 4.15; consideremos entonces el caso $c < 0$. Utilizando el Teorema 4.16 y el resultado para $c \geq 0$ obtenemos la Riemann-integrabilidad de $c \cdot f(x)$ y la relación

$$\int_X (c \cdot f(x)) dx = \int_X (-|c| \cdot f(x)) dx = - \int_X |c| \cdot f(x) dx = -|c| \int_X f(x) dx = c \cdot \int_X f(x) dx.$$

■

Teorema 4.21. Sean la funciones $f, g : \mathbb{R}^n \rightarrow \mathbb{R}$ Riemann-integrables sobre el conjunto X . Entonces también $f + g$ es Riemann-integrable sobre X , y se tiene que

$$\int_X (f(x) + g(x)) \, dx = \int_X f(x) \, dx + \int_X g(x) \, dx.$$

Demostración. El enunciado es una consecuencia inmediata del Teorema 4.14. ■

Teorema 4.22. Sean la funciones $f, g : \mathbb{R}^n \rightarrow \mathbb{R}$ Riemann-integrables sobre el conjunto X , y sea $f(x) \leq g(x)$ para todo $x \in X$. Entonces se tiene que

$$\int_X f(x) \, dx \leq \int_X g(x) \, dx.$$

Demostración. El enunciado es una consecuencia inmediata del Teorema 4.17. ■

Teorema 4.23. Sea la función $f : \mathbb{R}^n \rightarrow \mathbb{R}$ Riemann-integrable sobre los conjuntos X_1 y X_2 , donde $X_1 \cap X_2 = \emptyset$. Entonces f es también Riemann-integrable sobre $X_1 \cup X_2$, y se tiene que

$$\int_{X_1 \cup X_2} f(x) \, dx = \int_{X_1} f(x) \, dx + \int_{X_2} f(x) \, dx.$$

Demostración. El enunciado es una consecuencia inmediata del Teorema 4.18. ■

Definición 4.12. Sea $f : \mathbb{R}^n \rightarrow \mathbb{R}$ y $X \subset D(f)$. Para $x \in X$ definimos las funciones

$$f^+(x) := \begin{cases} f(x) & \text{si } f(x) \geq 0, \\ 0 & \text{si } f(x) < 0, \end{cases} \quad f^-(x) := \begin{cases} -f(x) & \text{si } f(x) < 0, \\ 0 & \text{si } f(x) \geq 0. \end{cases}$$

Evidentemente, se tiene lo siguiente:

$$\begin{aligned} f(x) &= f^+(x) - f^-(x); & |f(x)| &= f^+(x) + f^-(x) \quad \text{para todo } x \in X, \\ f_X(x) &= f_X^+(x) - f_X^-(x); & |f_X(x)| &= f_X^+(x) + f_X^-(x) \quad \text{para todo } x \in \mathbb{R}^n. \end{aligned}$$

Teorema 4.24. Sea la función $f : \mathbb{R}^n \rightarrow \mathbb{R}$ Riemann-integrable sobre el conjunto X . Entonces también f^+ y f^- son Riemann-integrables sobre X .

Demostración. Demostraremos la integrabilidad solamente para f^+ ; la demostración para f^- es análoga. Sea I un intervalo cerrado tal que $X \subset I$, y sea $P = \{I_1, \dots, I_m\}$ una partición arbitraria de I . Si definimos

$$M_k := \sup_{I_k} f_X(x), \quad m_k := \inf_{I_k} f_X(x), \quad M_k^+ := \sup_{I_k} f_X^+(x), \quad m_k^+ := \inf_{I_k} f_X^+(x),$$

obtenemos evidentemente

$$M_k^+ - m_k^+ \leq M_k - m_k$$

y obtenemos

$$0 \leq \bar{S}_P(f_X^+) - \underline{S}_P(f_X^+) = \sum_{k=1}^m (M_k^+ - m_k^+) \mu(I_k) \leq \sum_{k=1}^m (M_k - m_k) \mu(I_k) = \bar{S}_P(f_X) - \underline{S}_P(f_X).$$

Utilizando el criterio de integrabilidad de Riemann podemos concluir que f^+ es integrable sobre X . ■

Con la ayuda de este resultado obtenemos el siguiente teorema.

Teorema 4.25. *Sea la función $f : \mathbb{R}^n \rightarrow \mathbb{R}$ Riemann-integrable sobre el conjunto X . Entonces también $|f|$ es Riemann-integrable sobre X , y se tiene que*

$$\left| \int_X f(x) \, dx \right| \leq \int_X |f(x)| \, dx.$$

Demostración. La demostración es análoga a la demostración del resultado correspondiente del cálculo de las funciones de una variable. ■

Teorema 4.26. *Sean la funciones $f, g : \mathbb{R}^n \rightarrow \mathbb{R}$ Riemann-integrables sobre el conjunto X , Entonces también $f \cdot g$ es Riemann-integrable sobre X .*

Demostración. La demostración es análoga a la demostración del resultado correspondiente del cálculo de las funciones de una variable. ■

4.8. La medida de conjuntos

Consideremos ahora el problema de definir una medida, es decir, un contenido, para subconjuntos de $\mathbb{R}^n$. Para intervalos $I = [a, b] \subset \mathbb{R}^n$ ya hemos definido la medida n -dimensional

$$\mu(I) = \prod_{k=1}^n (b_k - a_k)$$

(ver Definición 4.1). Partiendo del concepto de la medida para intervalos, queremos definir la medida para subconjuntos $X \subset \mathbb{R}^n$ mucho más generales. Aproximaremos un conjunto X por uniones de números finitos de intervalos “desde afuera” y “desde el interior”.

Definición 4.13. *Sea $X \subset \mathbb{R}^n$ un conjunto acotado, y sea $I \subset \mathbb{R}^n$ un intervalo cerrado con $X \subset I$.*

1. *Para una partición P de I definimos*

$$\underline{M}_P(X) := \sum_{I_k \subset X} \mu(I_k), \quad \overline{M}_P(X) := \sum_{I_k \cap X \neq \emptyset} \mu(I_k).$$

2. *El número*

$$\underline{\mu}(X) := \sup_P \underline{M}_P(X)$$

se llama medida interior de Riemann de X .

3. *El número*

$$\overline{\mu}(X) := \inf_P \overline{M}_P(X)$$

se llama medida exterior de Riemann de X .

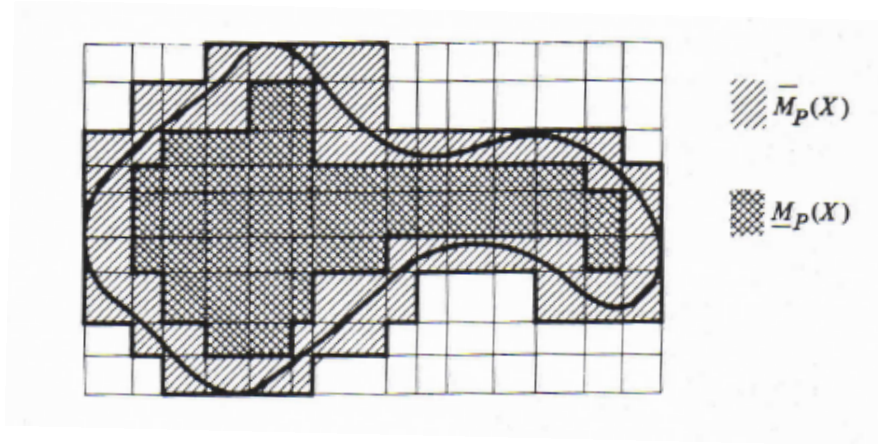


FIGURA 4.4. Ilustración de la Definición 4.13 para el caso $n = 2$. Las sumas $\underline{M}_P(X)$ y $\overline{M}_P(X)$ son aproximaciones del área de X por sumas de áreas de rectángulos desde el interior y desde afuera, respectivamente.

4. Se dice que el conjunto X es Riemann-medible si se tiene que

$$\underline{\mu}(X) = \overline{\mu}(X).$$

El valor común se llama medida n -dimensional de Riemann de X , denotada $\mu(X)$.

La Figura 4.4 ilustra la Definición 4.13 para el caso $n = 2$.

Es evidente que las cantidades $\underline{\mu}(X)$ y $\overline{\mu}(X)$ son independientes del intervalo particular considerado $I \subset \mathbb{R}^n$ con $X \subset I$. Por otro lado se tiene que

$$0 \leq \underline{\mu}(X) \leq \overline{\mu}(X),$$

y las Definiciones 4.1 y 4.13 son compatibles.

Mencionamos que en lugar de “medida n -dimensional de Riemann de X ” la cantidad $\mu(X)$ también se denomina por “contenido n -dimensional de Riemann” o “medida n -dimensional de Jordan”.

Ejemplo 4.4. Sea $X = [0, 1] \cap \mathbb{Q}$. Para una partición arbitraria P de $[0, 1]$ obtenemos $\underline{M}_P(X) = 0$ y $\overline{M}_P(X) = 1$, por lo tanto X no es Riemann-medible.

Ya vimos que existe una conexión entre la medida de un intervalo cerrado $I \subset \mathbb{R}^n$ y una integral, puesto que

$$\mu(I) = \int_I dx.$$

Demostraremos ahora una representación integral análoga para medidas generales (interiores y exteriores).

Teorema 4.27. Sea $X \subset \mathbb{R}^n$ un conjunto acotado y c_X su función característica.

1. Para las medidas interiores y exteriores de Riemann se tiene que

$$\underline{\mu}(X) = \int_{\underline{X}} c_X(x) dx = \int_{\underline{X}} dx,$$

$$\overline{\mu}(X) = \int_{\overline{X}} c_X(x) dx = \int_{\overline{X}} dx.$$

2. El conjunto X es Riemann-medible si y sólo si c_X es Riemann-integrable sobre X . En este caso se tiene que

$$\mu(X) = \int_X c_X(x) dx = \int_X dx.$$

Demostración. Sea $I \subset \mathbb{R}^n$ un intervalo cerrado con $X \subset I$.

1. Para una partición arbitraria $P = \{I_1, \dots, I_m\}$ de I se tiene que

$$\underline{S}_P(c_X) = \sum_{k=1}^m \inf_{I_k} c_X(x) \mu(I_k) = \sum_{I_k \subset X} \mu(I_k) = \underline{M}_P(X),$$

$$\overline{S}_P(c_X) = \sum_{k=1}^m \sup_{I_k} c_X(x) \mu(I_k) = \sum_{I_k \cap X \neq \emptyset} \mu(I_k) = \overline{M}_P(X),$$

lo que implica el primer enunciado.

2. El segundo enunciado es una consecuencia de la definición de la Riemann-medibilidad. ■

El Teorema 4.27 es muy importante dado que nos permite utilizar el cálculo integral para la determinación de la medida. En las demostraciones de los siguientes teoremas utilizaremos resultados ya conocidos para el cálculo integral.

Teorema 4.28. Sean $X_1 \subset \mathbb{R}^n$ y $X_2 \subset \mathbb{R}^n$ conjuntos Riemann-medibles con $X_1 \cap X_2 = \emptyset$. Entonces $X_1 \cup X_2$ es Riemann-medible, y se tiene que

$$\mu(X_1 \cup X_2) = \mu(X_1) + \mu(X_2).$$

Demostración. El Teorema 4.18 implica que

$$\underline{\mu}(X_1) + \underline{\mu}(X_2) \leq \underline{\mu}(X_1 \cup X_2) \leq \overline{\mu}(X_1 \cup X_2) \leq \overline{\mu}(X_1) + \overline{\mu}(X_2).$$

Pero $\underline{\mu}(X_1) = \overline{\mu}(X_1)$ y $\underline{\mu}(X_2) = \overline{\mu}(X_2)$, por lo tanto en esta cadena de desigualdades siempre tenemos igualdad. ■

Teorema 4.29. Sean $X_1 \subset \mathbb{R}^n$ y $X_2 \subset \mathbb{R}^n$ conjuntos Riemann-medibles con $X_1 \subset X_2$. Entonces el conjunto de diferencia $X_2 \setminus X_1$ es Riemann-medible, y se tiene que

$$\mu(X_2 \setminus X_1) = \mu(X_2) - \mu(X_1).$$

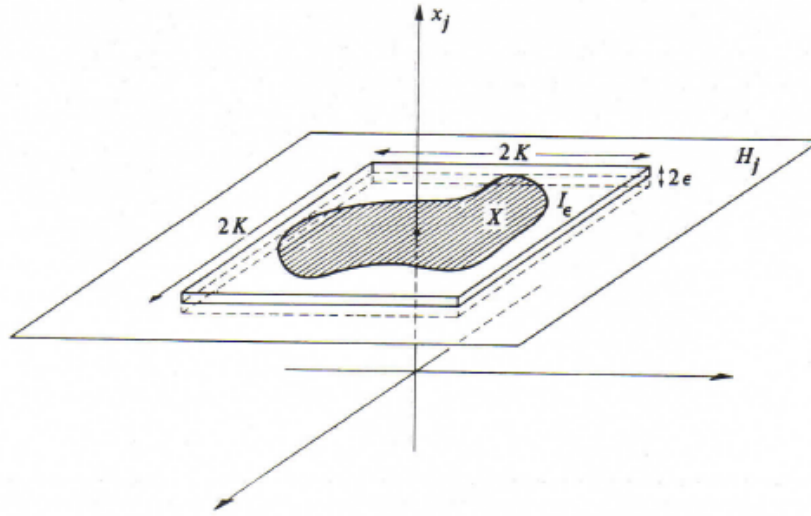


FIGURA 4.5. Ilustración de la demostración del Teorema 4.30.

Demostración. Consideremos los conjuntos X_1 y $X_2 \setminus X_1$. Se tiene que $X_1 \cap (X_2 \setminus X_1) = \emptyset$ y $X_1 \cup (X_2 \setminus X_1) = X_2$. La segunda y la tercera desigualdad del Teorema 4.18 nos entrega

$$\mu(X_2) \leq \underline{\mu}(X_2 \setminus X_1) + \mu(X_1) \leq \mu(X_2),$$

e intercambiando los conjuntos obtenemos de las mismas desigualdades que

$$\mu(X_2) \leq \mu(X_1) + \bar{\mu}(X_2 \setminus X_1) \leq \mu(X_2),$$

por lo tanto obtenemos

$$\underline{\mu}(X_2 \setminus X_1) = \bar{\mu}(X_2 \setminus X_1) = \mu(X_2 \setminus X_1), \quad \mu(X_2 \setminus X_1) = \mu(X_2) - \mu(X_1).$$

■

Definición 4.14. Un conjunto Riemann-medible X se llama conjunto nulo si $\mu(X) = 0$. Para el conjunto vacío $\emptyset$ definimos $\mu(\emptyset) = 0$.

Cada conjunto que contiene un número finito de puntos es, evidentemente, un conjunto nulo. De gran interés son aquellos conjuntos que pertenecen a un espacio de dimensión menor que n . El siguiente teorema muestra que estos conjuntos siempre son conjuntos nulos.

Teorema 4.30. Un conjunto $X \subset \mathbb{R}^n$ que está completamente contenido en un hiperplano

$$H_j = \{(x_1, \dots, x_n) \mid x_j = c = \text{const.}\}$$

es un conjunto nulo.

Demostración. Elegimos una constante $K > 0$ tal que

$$X \subset I = \{(x_1, \dots, x_n) \mid -K \leq x_i \leq K; i = 1, \dots, n\}.$$

En este caso, para cada $\varepsilon > 0$ el conjunto X está contenido en el intervalo n -dimensional

$$I_\varepsilon := \{(x_1, \dots, x_n) \mid -K \leq x_i \leq K (i \neq j); |x_j - c| < \varepsilon\},$$

y se tiene que

$$\bar{\mu}(X) \leq \bar{\mu}(I_\varepsilon) = \mu(I_\varepsilon) = (2K)^{n-1} \cdot 2\varepsilon$$

(ver Figura 4.5). Puesto que $\varepsilon > 0$ fue elegido arbitrario, se tiene que $\bar{\mu}(X) = 0$ y por lo tanto $\mu(X) = 0$. ■

Teorema 4.31. *Sea $X_1 \subset X_2$ y sea X_2 un conjunto nulo. Entonces también X_1 es un conjunto nulo.*

Demostración. Se tiene que $0 \leq \bar{\mu}(X_1) \leq \bar{\mu}(X_2) = 0$, por lo tanto $\mu(X_1) = 0$. ■

Definición 4.15. *Sea $X \subset \mathbb{R}^n$.*

1. *El conjunto de los puntos interiores de X se llama el interior de X , denotado X^0 .*
2. *Un punto $x_0 \in \mathbb{R}^n$ se llama punto de frontera de X si en cada vecindad de x_0 existen por lo menos un punto de X y un punto de $\mathbb{R}^n \setminus X$. El conjunto de todos los puntos de frontera de X se llama frontera o borde de X , denotado ∂X .*
3. *El conjunto $\bar{X} := X \cup \partial X$ se llama clausura de X .*

Teorema 4.32. *Un conjunto acotado $X \subset \mathbb{R}^n$ es Riemann-medible si y sólo si $\mu(\partial X) = 0$.*

Demostración. Sea $I \subset \mathbb{R}^n$ un intervalo cerrado tal que $X \subset I$.

1. Sea $\underline{\mu}(\partial X) = 0$. Entonces para $\varepsilon > 0$ existe una partición $P = \{I_1, \dots, I_m\}$ de I tal que $\bar{M}_P(\partial X) < \varepsilon$. Ahora se tiene que

$$\begin{aligned} \bar{\mu}(X) - \underline{\mu}(X) &\leq \bar{M}_P(X) - \underline{M}_P(X) \\ &= \sum_{I_k \cap X \neq \emptyset} \mu(I_k) - \sum_{I_k \subset X} \mu(I_k) \\ &\leq \sum_{I_k \cap \partial X \neq \emptyset} \mu(I_k) = \bar{M}_P(\partial X) < \varepsilon. \end{aligned}$$

Puesto que $\varepsilon > 0$ fue elegido arbitrario, se tiene que $\bar{\mu}(X) = \underline{\mu}(X)$, por lo tanto X es Riemann-medible.

2. Sea X Riemann-medible. Entonces para cada $\varepsilon > 0$ existe una partición P de I tal que

$$\bar{M}_P(X) - \underline{M}_P(X) < \varepsilon.$$

Ahora podemos representar la frontera ∂X como $\partial X = X_1 \cup X_2$, donde

$$X_1 = \left\{ x \mid x \in \partial X, x \notin \bigcup_{I_k \subset X} I_k \right\}, \quad X_2 = \partial X \setminus X_1.$$

Los puntos de X_2 pertenecen a la frontera de ciertos intervalos que entregan contribuciones a $\underline{M}_P(X)$, es decir pertenecen a la unión de un número finito de hiperplanos. Según el Teorema 4.30, entonces se tiene que $\mu(X_2) = 0$, además se tiene que

$$\bar{\mu}(X_1) \leq \bar{M}_P(X_1) = \bar{M}_P(X) - \underline{M}_P(X) < \varepsilon,$$

luego, según el Teorema 4.18,

$$\bar{\mu}(\partial X) \leq \bar{\mu}(X_1) + \bar{\mu}(X_2) = \bar{\mu}(X_1) < \varepsilon.$$

Puesto que $\varepsilon > 0$ es arbitrario, se tiene que $\bar{\mu}(\partial X) = 0$, por lo tanto también $\mu(\partial X) = 0$. ■

Con la ayuda de este teorema se puede demostrar el siguiente teorema.

Teorema 4.33. *Si $X \subset \mathbb{R}^n$ es un conjunto Riemann-medible, entonces también X^0 y $\bar{X}$ son Riemann-medibles, y se tiene que*

$$\mu(X) = \mu(X^0) = \mu(\bar{X}).$$

Demostración. Tarea. ■

Teorema 4.34. *Sean $X_1 \subset \mathbb{R}^n$ y $X_2 \subset \mathbb{R}^n$ conjuntos Riemann-medibles. Entonces también $X_1 \cup X_2$ y $X_1 \cap X_2$ son Riemann-medibles, y se tiene que*

$$\mu(X_1 \cup X_2) + \mu(X_1 \cap X_2) = \mu(X_1) + \mu(X_2).$$

Demostración.

1. Se tiene que

$$\partial(X_1 \cap X_2) \subset \partial X_1 \cup \partial X_2,$$

por lo tanto $\partial(X_1 \cap X_2)$ también es un conjunto nulo, y según el Teorema 4.32 el conjunto $X_1 \cap X_2$ es Riemann-medible.

2. Se tiene que

$$X_1 \cup X_2 = X_1 \cup (X_2 \setminus (X_1 \cap X_2)), \quad X_1 \cap (X_2 \setminus (X_1 \cap X_2)) = \emptyset,$$

entonces de la Riemann-medibilidad de $X_2 \setminus (X_1 \cap X_2)$ sigue según el Teorema 4.28 la Riemann-medibilidad de $X_1 \cup X_2$, y obtenemos que

$$\begin{aligned} \mu(X_1 \cup X_2) &= \mu(X_1) + \mu(X_2 \setminus (X_1 \cap X_2)) \\ &= \mu(X_1) + \mu(X_2) - \mu(X_1 \cap X_2). \end{aligned}$$
■

Evidentemente, una versión análoga de este teorema es válida para números finitos de conjuntos. En particular, si $X_1 \subset \mathbb{R}^n, \dots, X_N \subset \mathbb{R}^n$ son conjuntos Riemann-medibles, entonces también

$$X = \bigcup_{\nu=1}^N X_\nu$$

es Riemann-medible, y se tiene que

$$\mu(X) \leq \sum_{\nu=1}^N \mu(X_\nu).$$

Obtenemos la siguiente caracterización de conjuntos nulos que a veces es útil.

Teorema 4.35. *Un conjunto $X \subset \mathbb{R}^n$ es un conjunto nulo si y sólo si para cada $\varepsilon > 0$ existe un número finito de intervalos cerrados $I_1, \dots, I_N$ tales que*

$$X \subset \bigcup_{\nu=1}^N I_\nu \quad y \quad \sum_{\nu=1}^N \mu(I_\nu) < \varepsilon.$$

Demostración.

1. Supongamos que para $\varepsilon > 0$ existe un número finito de intervalos cerrados $I_1, \dots, I_N$ tales que

$$X \subset J = \bigcup_{\nu=1}^N I_\nu \quad y \quad \sum_{\nu=1}^N \mu(I_\nu) < \varepsilon.$$

Entonces el conjunto J es Riemann-medible, y se tiene que

$$\bar{\mu}(X) \leq \mu(J) \leq \sum_{\nu=1}^N \mu(I_\nu) < \varepsilon.$$

Puesto que $\varepsilon > 0$ fue elegido arbitrario, se tiene que $\mu(X) = 0$.

2. Sea $\mu(X) = 0$. Sea $I \subset \mathbb{R}^n$ un intervalo cerrado tal que $X \subset I$. Entonces para $\varepsilon > 0$ existe una partición $\{I_1, \dots, I_m\}$ de I tal que

$$\bar{M}_P(X) = \sum_{I_k \cap X \neq \emptyset} \mu(I_k) < \varepsilon.$$

En este caso, el número finito de intervalos cerrados I_k con $I_k \cap X \neq \emptyset$ tiene la propiedad deseada. ■

4.9. Conjuntos nulos especiales

El siguiente teorema es importante en las aplicaciones porque nos permite en muchos casos decidir la Riemann-medibilidad de conjuntos.

Teorema 4.36. *Sea la función $f = (f_1, \dots, f_m) : \mathbb{R}^n \rightarrow \mathbb{R}^m$ continua sobre el conjunto compacto $X \subset \mathbb{R}^n$. Entonces el siguiente conjunto es un conjunto nulo de $\mathbb{R}^{n+m}$:*

$$Z := \left\{ x \in \mathbb{R}^{n+m} \mid x = (x_1, \dots, x_n, f_1(x_1, \dots, x_n), \dots, f_m(x_1, \dots, x_n)), (x_1, \dots, x_n) \in X \right\}.$$

Demostración. Sea $I \subset \mathbb{R}^n$ un intervalo cerrado tal que $X \subset I$, además sea $\varepsilon > 0$ dado. Puesto que f es uniformemente continua sobre X existe una partición $P = \{I_1, \dots, I_l\}$ de I tal que para $i = 1, \dots, m$ se tiene que

$$\forall x', x'' \in I_k \cap X \quad (k = 1, \dots, l) : \quad |f_i(x') - f_i(x'')| < \varepsilon.$$

Entonces, aquellos puntos en Z para los cuales se tiene que $x = (x_1, \dots, x_n) \in I_k \cap X$ pertenecen a un intervalo cerrado $(n+m)$ -dimensional de la medida $(2\varepsilon)^m \mu(I_k)$, por lo tanto Z está

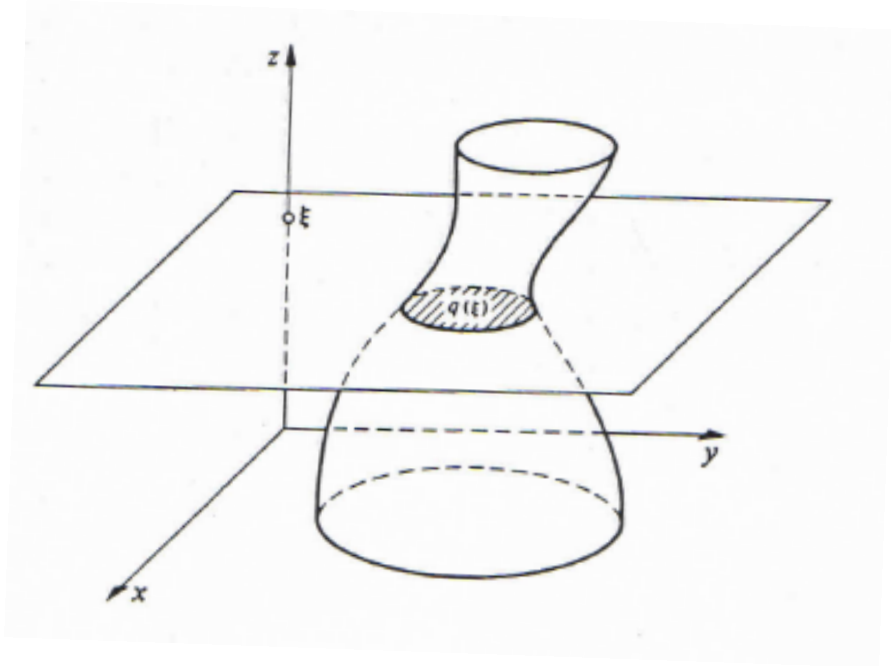


FIGURA 4.6. Ilustración del Principio de Cavalieri.

contenido en la unión de un número finito de intervalos cerrados de $\mathbb{R}^{n+m}$ cuya medida total es

$$\sum_{k=1}^l (2\varepsilon)^m \mu(I_k) = (2\varepsilon)^m \mu(I).$$

Ahora la afirmación es una consecuencia del Teorema 4.35. ■

Ejemplo 4.5. Sea $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ y sea f continua sobre $[a, b]$. En este caso el conjunto de puntos

$$C_f := \{(x, y) \in \mathbb{R}^2 \mid y = f(x), x \in [a, b]\}$$

es un conjunto nulo de $\mathbb{R}^2$.

Ejemplo 4.6. Sea $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ y sea f continua sobre el intervalo cerrado I . Entonces el conjunto de puntos

$$F_f := \{(x, y, z) \in \mathbb{R}^3 \mid z = f(x, y), (x, y) \in I\}$$

es un conjunto nulo de $\mathbb{R}^3$.

4.10. El principio de Cavalieri

Teorema 4.37 (Cavalieri). Sea $S \subset \mathbb{R}^n$ un conjunto Riemann-medible y

$$S \subset I := \{(x_1, \dots, x_n) \mid a_i \leq x_i \leq b_i, i = 1, \dots, n\}.$$

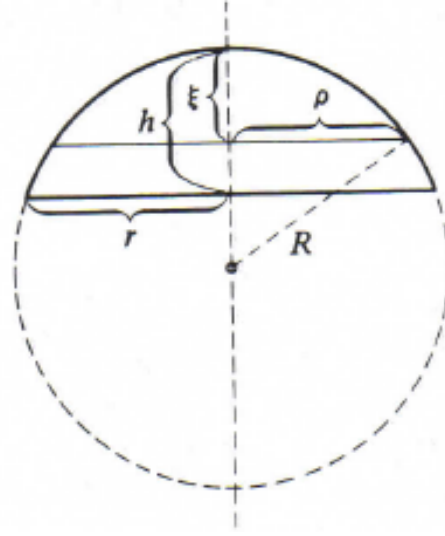


FIGURA 4.7. Ilustración del Ejemplo 4.7.

Ahora, si para un índice ν fijo ($1 \leq \nu \leq n$) el conjunto

$$S_{\xi}^{(\nu)} = S \cap \{(x_1, \dots, x_n) \mid x_{\nu} = \xi\}$$

es Riemann-medible para todo $\xi \in [a_{\nu}, b_{\nu}]$ en el espacio $(n-1)$ -dimensional de las coordenadas $(x_1, \dots, x_{\nu-1}, x_{\nu+1}, \dots, x_n)$ con la medida $q_{\nu}(\xi)$, entonces se tiene que

$$\mu(S) = \int_{a_{\nu}}^{b_{\nu}} q_{\nu}(\xi) d\xi$$

(ver Figura 4.6).

Demostración. Según el Teorema 4.11 se tiene que

$$\begin{aligned} \mu(S) &= \int_S dx = \int_I c_S(x) dx = \int_{a_{\nu}}^{b_{\nu}} \left[\int_{I_{x_{\nu}}} c_S(x) d(x_1, \dots, x_{\nu-1}, x_{\nu+1}, \dots, x_n) \right] dx_{\nu} \\ &= \int_{a_{\nu}}^{b_{\nu}} \left[\int_{S_{\xi}^{(\nu)}} d(x_1, \dots, x_{\nu-1}, x_{\nu+1}, \dots, x_n) \right] d\xi. \end{aligned}$$

En virtud de

$$q_{\nu}(\xi) = \int_{S_{\xi}^{(\nu)}} d(x_1, \dots, x_{\nu-1}, x_{\nu+1}, \dots, x_n)$$

se llega a la afirmación. ■

Ejemplo 4.7. Se desea determinar el volumen V del segmento de una bola de la altura h y del radio r (ver Figura 4.7). Aquí se tiene que

$$q(\xi) = \pi \varrho^2, \quad \text{donde} \quad \varrho^2 = R^2 - (R - \xi)^2 = 2R\xi - \xi^2 \quad (0 \leq \xi \leq h).$$

Según el principio de Cavalieri obtenemos

$$V = \int_0^h q(\xi) d\xi = \pi \int_0^h (2R\xi - \xi^2) d\xi = \pi \left(Rh^2 - \frac{h^3}{3} \right).$$

En virtud de $Rh = (h^2 + r^2)/2$ obtenemos

$$V = \pi \left(\frac{h^3 + r^2 h}{2} - \frac{h^3}{3} \right) = \frac{\pi h}{6} (3r^2 + h^2).$$

4.11. Los teoremas de valores intermedios del cálculo integral n -dimensional

Los siguientes teoremas se utilizan principalmente para acotar integrales. Aquí se utiliza esencialmente el concepto de la Riemann-integrabilidad.

Teorema 4.38. Sea $X \subset \mathbb{R}^n$ un conjunto Riemann-medible. Sea la función $f : \mathbb{R}^n \rightarrow \mathbb{R}$ acotada sobre X , y como siempre sea

$$m(f) = \inf_X f(x), \quad M(f) = \sup_X f(x).$$

Entonces se tiene que

$$m(f)\mu(X) \leq \int_{\bar{X}} f(x) dx \leq \int_X f(x) dx \leq M(f)\mu(X).$$

Demostración. Sobre X se tiene que

$$m(f) \leq f(x) \leq M(f),$$

por lo tanto

$$\begin{aligned} m(f)\mu(X) &= \int_{\bar{X}} m(f) dx \leq \int_{\bar{X}} f(x) dx \leq \int_X f(x) dx \\ &\leq \int_X M(f) dx = M(f)\mu(X). \end{aligned}$$

■

El siguiente teorema expresa una consecuencia inmediata.

Teorema 4.39 (Primer Teorema del Valor Intermedio). Sea $X \subset \mathbb{R}^n$ un conjunto Riemann-medible y sea f Riemann-integrable sobre X .

1. En este caso existe un η tal que $m(f) \leq \eta \leq M(f)$ que satisface

$$\int_X f(x) dx = \eta \cdot \mu(X).$$

2. Si, además, X es compacto y conexo y la función f es continua sobre X , entonces existe un $\xi \in X$ tal que

$$\int_X f(x) \, dx = f(\xi) \cdot \mu(X). \quad (4.6)$$

Demostración.

1. El primer enunciado es una consecuencia del Teorema 4.38 y de la Definición 4.11.
2. Sean satisfechas las presuposiciones adicionales de la segunda afirmación. Si f es constante, cualquier $\xi \in X$ satisface (4.6). Si f no es constante, y por lo tanto $m(f) < M(f)$, se tiene según el Teorema 1.18 que según el Teorema 1.17, la imagen $f(X)$ es el intervalo cerrado $[m(f), M(f)]$, por lo tanto existe un $\xi \in X$ tal que $f(\xi) = \eta$. Esto concluye la demostración de la segunda afirmación. ■

Teorema 4.40 (Primer Teorema del Valor Intermedio Extendido). *Sea $X \subset \mathbb{R}^n$ un conjunto acotado y sean f y g Riemann-integrables sobre X , y sea $g(x) \geq 0$ para todo $x \in X$.*

1. En este caso existe un η tal que $m(f) \leq \eta \leq M(f)$ que satisface

$$\int_X f(x)g(x) \, dx = \eta \cdot \int_X g(x) \, dx.$$

2. Si, además, X es compacto y conexo y la función f es continua sobre X , entonces existe un $\xi \in X$ tal que

$$\int_X f(x)g(x) \, dx = f(\xi) \cdot \int_X g(x) \, dx. \quad (4.7)$$

Demostración. Puesto que $g(x) \geq 0$ se tiene que

$$m(f)g(x) \leq f(x)g(x) \leq M(f)g(x) \quad \text{para todo } x \in X,$$

entonces según el Teorema 4.22

$$m(f) \cdot \int_X g(x) \, dx \leq \int_X f(x)g(x) \, dx \leq M(f) \cdot \int_X g(x) \, dx.$$

Ahora, si

$$\int_X g(x) \, dx = 0,$$

el primer enunciado es válido para todo η y el segundo es válido para todo $\xi \in X$. Si

$$\int_X g(x) \, dx > 0,$$

podemos elegir

$$\eta = \frac{\int_X f(x)g(x) \, dx}{\int_X g(x) \, dx},$$

y resulta el primer enunciado. Si X es compacto y conexo y f es continua sobre X , entonces tal como en la demostración del Teorema 4.39 existe un $\xi \in X$ tal que $f(\xi) = \eta$, y llegamos al segundo enunciado. ■

Ejemplo 4.8. Sea $X_0 \subset \mathbb{R}^n$ un conjunto abierto, y sea la función $f : \mathbb{R}^n \rightarrow \mathbb{R}$ continua sobre X_0 . Supongamos que para cada conjunto medible $X \subset X_0$ se tiene que

$$\int_X f(x) \, dx = 0.$$

En este caso podemos concluir que $f(x) \equiv 0$ sobre X_0 , puesto que si fuera $f(x) \not\equiv 0$, existiría un $x^0 \in X_0$ tal que $f(x^0) \neq 0$; sin pérdida de la generalidad supongamos que $f(x^0) > 0$. En este caso existe un $\delta > 0$ tal que

$$Y = \{x \mid d(x, x^0) \leq \delta\} \subset X_0, \quad f(x) > 0 \quad \text{para todo } x \in Y.$$

El conjunto Y es compacto, conexo y medible con $\mu(Y) > 0$. Entonces, según el Teorema 4.39 existe un $\xi \in Y$ tal que

$$\int_Y f(x) \, dx = f(\xi) \cdot \mu(Y) > 0,$$

lo que contradice la hipótesis.

4.12. La integrabilidad sobre conjuntos generales

Teorema 4.41. Sea $X \subset \mathbb{R}^n$ un conjunto Riemann-medible, y sea f continua sobre $\bar{X}$. Entonces f es Riemann-integrable sobre X .

Demostración. Sea $|f(x)| \leq M$ sobre X y sea I un intervalo cerrado con $I \supset X$. Puesto que $\mu(\partial X) = 0$, podemos elegir una partición $P = \{I_1, \dots, I_m\}$ de I de tal modo que

$$\sum_{I_k \cap \partial X \neq \emptyset} \mu(I_k) < \varepsilon.$$

Puesto que f es uniformemente continua sobre $\bar{X}$ podemos elegir los diámetros de los demás I_k tan pequeños que sobre cada intervalo I_k se satisface que

$$\sup_{I_k} f_X(x) - \inf_{I_k} f_X(x) < \varepsilon.$$

Así obtenemos que

$$\bar{S}_P(f_X) - \underline{S}_P(f_X) \leq \varepsilon \sum_{I_k \cap \partial X = \emptyset} \mu(I_k) + 2M \sum_{I_k \cap \partial X \neq \emptyset} \mu(I_k) \leq \varepsilon \cdot \mu(I) + 2M\varepsilon,$$

lo cual implica el enunciado del teorema en virtud del Teorema 4.19. ■

Teorema 4.42. *Sea $X_0 \subset \mathbb{R}^n$ un conjunto acotado, y sea la función $f : \mathbb{R}^n \rightarrow \mathbb{R}$ Riemann-integrable sobre X , además sea X_0 un conjunto nulo tal que $X_0 \subset X$. Entonces cada función g acotada sobre X que satisface*

$$g(x) = f(x) \quad \text{para } x \in X \setminus X_0$$

es Riemann-integrable sobre X , y se tiene que

$$\int_X f(x) \, dx = \int_X g(x) \, dx.$$

Demostración. Puesto que $\mu(X_0) = 0$ sabemos que según el Teorema 4.38 que la función $h = f - g$ es Riemann-integrable sobre X_0 y que se tiene que

$$\int_{X_0} h(x) \, dx = 0.$$

Sobre $X \setminus X_0$ se tiene que $h = 0$, por lo tanto h es igualmente Riemann-integrable sobre $X \setminus X_0$. Ahora, según el Teorema 4.23,

$$\int_X h(x) \, dx = \int_{X \setminus X_0} h(x) \, dx + \int_{X_0} h(x) \, dx = \int_{X \setminus X_0} h(x) \, dx = 0,$$

lo que concluye la demostración del teorema. ■

Capítulo 5

La computación de integrales

5.1. Dominios normales

Definición 5.1. Sea S_{x_ν} un subconjunto del $(x_1, \dots, x_{\nu-1}, x_{\nu+1}, \dots, x_n)$ -espacio $(n-1)$ -dimensional. Para los puntos de S_{x_ν} escribimos

$$x^\nu = (x_1, \dots, x_{\nu-1}, x_{\nu+1}, \dots, x_n).$$

Definición 5.2. Sea $S \subset \mathbb{R}^n$ un conjunto acotado.

1. Se dice que S es proyectable en la dirección del eje x_ν o brevemente x_ν -proyectable si en el $(x_1, \dots, x_{\nu-1}, x_{\nu+1}, \dots, x_n)$ -espacio (de la dimensión $(n-1)$) existen un conjunto medible y cerrado S_{x_ν} y dos funciones continuas $\underline{x}_\nu$ y $\bar{x}_\nu$ tales que

$$\bar{S} = \bigcup_{x^\nu \in S_{x_\nu}} \{(x_1, \dots, x_n) \mid \underline{x}_\nu(x^\nu) \leq x_\nu \leq \bar{x}_\nu(x^\nu)\}.$$

2. Se dice que S es un recinto estándar si para cada $\nu = 1, \dots, n$, S es proyectable en la dirección del eje x_ν .

Comentamos que en el caso $n = 2$, según esta definición, un conjunto $S \subset \mathbb{R}^2$ es proyectable

1. en la dirección del eje x si existen dos funciones $\underline{x}(y)$ y $\bar{x}(y)$ continuas sobre un subconjunto S_y medible y cerrado del eje y tales que

$$\bar{S} = \bigcup_{y \in S_y} \{(x, y) \mid \underline{x}(y) \leq x \leq \bar{x}(y)\},$$

2. en la dirección del eje y si existen dos funciones $\underline{y}(x)$ y $\bar{y}(x)$ continuas sobre un subconjunto S_x medible y cerrado del eje x tales que

$$\bar{S} = \bigcup_{x \in S_x} \{(x, y) \mid \underline{y}(x) \leq y \leq \bar{y}(x)\}.$$

La Figura 5.1 muestra algunos ejemplos.

En el caso $n = 3$, un conjunto $S \subset \mathbb{R}^3$ es proyectable, por ejemplo, en la dirección del eje z si existen dos funciones $\underline{z}(x, y)$ y $\bar{z}(x, y)$ continuas sobre un subconjunto S_z medible y cerrado del plano (x, y) tales que

$$\bar{S} = \bigcup_{(x, y) \in S_z} \{(x, y, z) \mid \underline{z}(x, y) \leq z \leq \bar{z}(x, y)\},$$

ver Figura 5.2.

Teorema 5.1. Sea la función $f(x, y)$ Riemann-integrable sobre el conjunto $S \subset \mathbb{R}^2$.

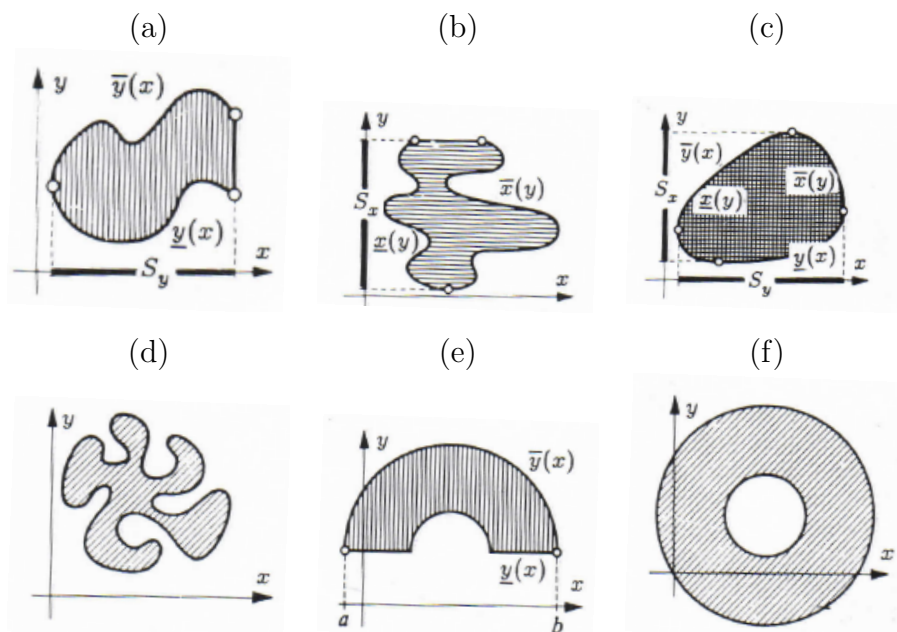


FIGURA 5.1. Ilustración de la Definición 5.2 para $n = 2$. El conjunto dibujado es (a) y -proyectable, pero no x -proyectable, (b) x -proyectable, pero no y -proyectable, (c) x -proyectable e y -proyectable (recinto estándar), (d) no x -proyectable ni y -proyectable, (e) y -proyectable, pero no x -proyectable, (f) no x -proyectable ni y -proyectable.

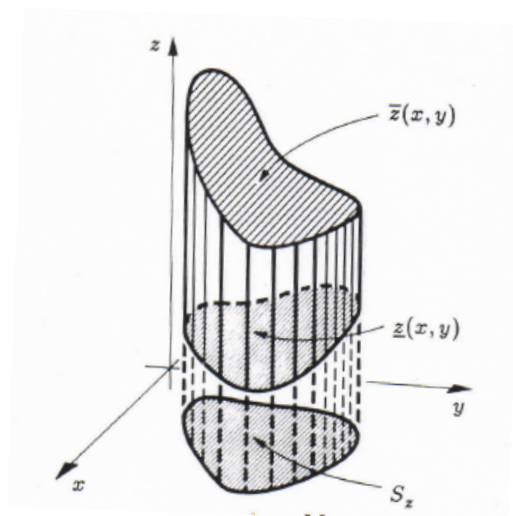


FIGURA 5.2. Ilustración de la Definición 5.2 para $n = 3$.

1. Sea S proyectable en la dirección del eje y . Si para cada $x \in S_y$ existe la integral

$$\int_{\underline{y}(x)}^{\bar{y}(x)} f(x, y) dy,$$

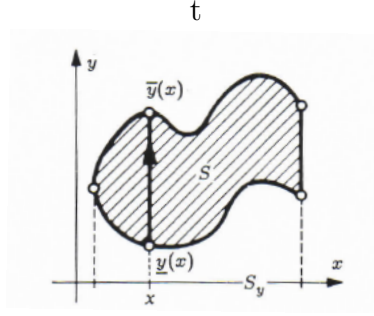


FIGURA 5.3. Ilustración de la demostración del Teorema 5.1.

entonces también existe la integral iterada

$$\int_{S_y} \left[\int_{\underline{y}(x)}^{\bar{y}(x)} f(x, y) dy \right] dx,$$

y se tiene que

$$\iint_S f(x, y) d(x, y) = \int_{S_y} \left[\int_{\underline{y}(x)}^{\bar{y}(x)} f(x, y) dy \right] dx.$$

La Figura 5.3 ilustra este caso.

2. Sea S proyectable en la dirección del eje x . Si para cada $y \in S_x$ existe la integral

$$\int_{\underline{x}(y)}^{\bar{x}(y)} f(x, y) dx,$$

entonces también existe la integral iterada

$$\int_{S_x} \left[\int_{\underline{x}(y)}^{\bar{x}(y)} f(x, y) dx \right] dy,$$

y se tiene que

$$\iint_S f(x, y) d(x, y) = \int_{S_x} \left[\int_{\underline{x}(y)}^{\bar{x}(y)} f(x, y) dx \right] dy.$$

Demostración. Demostraremos solamente la primera afirmación; la demostración de la segunda es análoga. Para tal efecto consideremos un intervalo

$$I = \{(x, y) \mid A \leq x \leq B, C \leq y \leq D\}$$

tal que $S \subset I$, entonces se tiene que

$$\iint_S f(x, y) d(x, y) = \iint_I f_S(x, y) d(x, y).$$

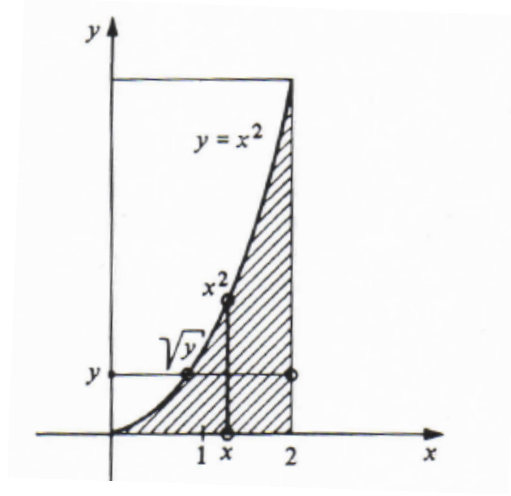


FIGURA 5.4. Ilustración del Ejemplo 5.1.

Para cada $x \in S_y$ fijo obtenemos

$$\int_{\underline{y}(x)}^{\bar{y}(x)} f(x, y) \, dy = \int_C^D f_S(x, y) \, dy.$$

En virtud del Teorema 4.10 obtenemos ahora

$$\iint_I f_S(x, y) \, d(x, y) = \int_A^B \left[\int_C^D f_S(x, y) \, dy \right] dx = \int_{S_y} \left[\int_{\underline{y}(x)}^{\bar{y}(x)} f(x, y) \, dy \right] dx,$$

lo que concluye la demostración. ■

Ejemplo 5.1. *Sea*

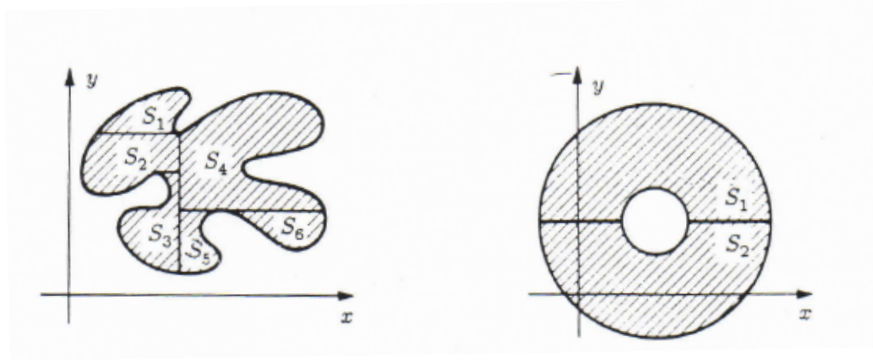
$$S = \{(x, y) \mid 0 \leq x \leq 2, 0 \leq y \leq x^2\},$$

y consideremos la función $f(x, y) = x^2 + y^2$ definida sobre S (ver Figura 5.4). Entonces S es proyectable en la dirección del eje y , y según el Teorema 5.1 se tiene que

$$\iint_S f(x, y) \, d(x, y) = \iint_S (x^2 + y^2) \, d(x, y) = \int_0^2 \left[\int_0^{x^2} (x^2 + y^2) \, dy \right] dx.$$

Aquí obtenemos

$$\begin{aligned} \int_0^{x^2} (x^2 + y^2) \, dy &= \left[x^2 y + \frac{y^3}{3} \right]_{y=0}^{y=x^2} = x^4 + \frac{x^6}{3}, \\ \int_0^2 \left(x^4 + \frac{x^6}{3} \right) dx &= \left[\frac{x^5}{5} + \frac{x^7}{21} \right]_{x=0}^{x=2} = \frac{32}{5} + \frac{128}{21} = \frac{1312}{105}, \end{aligned}$$

FIGURA 5.5. Descomposición de S en conjuntos $S_1, \dots, S_N$ proyectables.

por lo tanto

$$\iint_S f(x, y) \, d(x, y) = \frac{1312}{105}.$$

Puesto que S también es proyectable en la dirección del eje x , también se tiene que

$$\begin{aligned} \iint_S f(x, y) \, d(x, y) &= \iint_S (x^2 + y^2) \, d(x, y) \\ &= \int_0^4 \left[\int_{\sqrt{y}}^2 (x^2 + y^2) \, dx \right] dy \\ &= \int_0^4 \left(\frac{8}{3} + 2y^2 - \frac{1}{3}y^{3/2} - y^{5/2} \right) dy \\ &= \left[\frac{8}{3}y + \frac{2}{3}y^3 - \frac{2}{15}y^{5/2} - \frac{2}{7}y^{7/2} \right]_{y=0}^{y=4} = \frac{1312}{105}. \end{aligned}$$

El Teorema 5.1 nos permite la computación de integrales de funciones continuas sobre conjuntos proyectables. Pero si un conjunto $S \subset \mathbb{R}^2$ no es proyectable, posiblemente se puede descomponer S en un número finito de subconjuntos $S_1, \dots, S_N$ proyectables (ver Figura 5.5). En este caso obtenemos

$$\iint_S f(x, y) \, d(x, y) = \iint_{S_1} f(x, y) \, d(x, y) + \dots + \iint_{S_N} f(x, y) \, d(x, y),$$

donde cada una de las integrales

$$\iint_{S_i} f(x, y) \, d(x, y), \quad i = 1, \dots, N$$

puede ser calculada como integral iterada.

El siguiente teorema representa una generalización del Teorema 4.11 que nos permite la computación de integrales sobre conjuntos n -dimensionales.

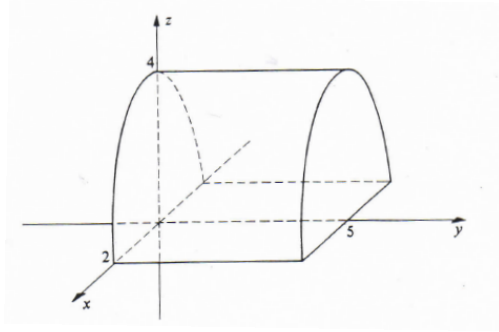


FIGURA 5.6. Ilustración del Ejemplo 5.2.

Teorema 5.2. Sea la función $f : \mathbb{R}^n \rightarrow \mathbb{R}$ Riemann-integrable sobre el conjunto $S \subset \mathbb{R}^n$. Sea S proyectable en la dirección del eje x_ν . Si para cada $x^\nu = (x_1, \dots, x_{\nu-1}, x_{\nu+1}, \dots, x_n) \in S_{x_\nu}$ existe la integral

$$\int_{\underline{x}_\nu(x^\nu)}^{\bar{x}_\nu(x^\nu)} f(x_1, \dots, x_\nu, \dots, x_n) dx_\nu,$$

entonces también existe la integral iterada

$$\int_{S_{x_\nu}} \left[\int_{\underline{x}_\nu(x^\nu)}^{\bar{x}_\nu(x^\nu)} f(x_1, \dots, x_\nu, \dots, x_n) dx_\nu \right] d(x_1, \dots, x_{\nu-1}, x_{\nu+1}, \dots, x_n),$$

y su valor es

$$\int_S f(x) dx.$$

Demostración. La demostración es análoga a la del Teorema 5.1 mediante la aplicación del Teorema 4.11. ■

Para el caso $n = 3$ obtenemos para un conjunto S , que sea proyectable (por ejemplo) en la dirección del eje z , la siguiente fórmula:

$$\iiint_S f(x, y, z) d(x, y, z) = \iint_{S_z} \left[\int_{\underline{z}(x,y)}^{\bar{z}(x,y)} f(x, y, z) dz \right] d(x, y).$$

Si, además, el conjunto S_z es proyectable en la dirección del eje y , obtenemos

$$\iiint_S f(x, y, z) d(x, y, z) = \int_{S_{zy}} \left\{ \int_{\underline{y}(x)}^{\bar{y}(x)} \left[\int_{\underline{z}(x,y)}^{\bar{z}(x,y)} f(x, y, z) dz \right] dy \right\} dx.$$

Se obtienen fórmulas análogas si S es proyectable en otra dirección.

Ejemplo 5.2. Se desea calcular el volumen V del conjunto

$$S = \{(x, y, z) \mid -2 \leq x \leq 2, 0 \leq y \leq 5, 0 \leq z \leq 4 - x^2\}$$

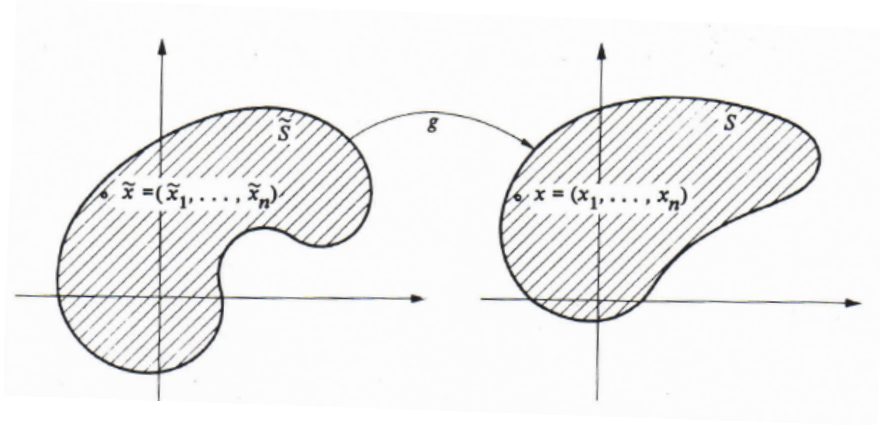


FIGURA 5.7. Ilustración de la Definición 5.3.

(ver Figura 5.6). Si definimos

$$S_z = \{(x, y) \mid -2 \leq x \leq 2, 0 \leq y \leq 5\},$$

obtenemos según el Teorema 5.2

$$\begin{aligned} V &= \iiint_S d(x, y, z) \\ &= \iint_{S_z} \left[\int_0^{4-x^2} dz \right] d(x, y) = \iint_{S_z} (4 - x^2) d(x, y) \\ &= \int_{-2}^2 \left[\int_0^5 (4 - x^2) dy \right] dx = \int_{-2}^2 (20 - 5x^2) dx = \frac{160}{3}. \end{aligned}$$

5.2. La regla de substitución para integrales n -dimensionales

Durante la computación de integrales n -dimensionales puede aparecer la necesidad de integrar una función sobre un subconjunto de $\mathbb{R}^n$ cuya descripción por coordenadas cartesianas puede ser muy complicada. En tales casos se definen coordenadas nuevas.

Consideremos primeramente el espacio $R = \mathbb{R}^n$ de los puntos $x = (x_1, \dots, x_n)$, $x_i \in \mathbb{R}$, $i = 1, \dots, n$. En muchas situaciones podemos describir un subconjunto $S \subset R$ por un conjunto $\tilde{S}$ de puntos $\tilde{x} = (\tilde{x}_1, \dots, \tilde{x}_n)$ que son elementos de un segundo espacio $\tilde{R} = \mathbb{R}^n$. Esto se realiza asignando de manera biyectiva un punto $\tilde{x} = (\tilde{x}_1, \dots, \tilde{x}_n) \in \tilde{S}$ a cada punto $x = (x_1, \dots, x_n) \in S$. Se dice que el punto $x = (x_1, \dots, x_n)$ *posee las coordenadas nuevas* $\tilde{x} = (\tilde{x}_1, \dots, \tilde{x}_n)$, ver Figura 5.7.

En virtud de lo anterior podemos introducir coordenadas nuevas mediante una aplicación biyectiva $g = (g_1, \dots, g_n) : \tilde{R} = \mathbb{R}^n \rightarrow \mathbb{R}^n$ con $D(g) = \tilde{S}$, donde

$$g : \begin{cases} x_1 &= g_1(\tilde{x}_1, \dots, \tilde{x}_n), \\ &\vdots \\ x_n &= g_n(\tilde{x}_1, \dots, \tilde{x}_n), \end{cases}$$

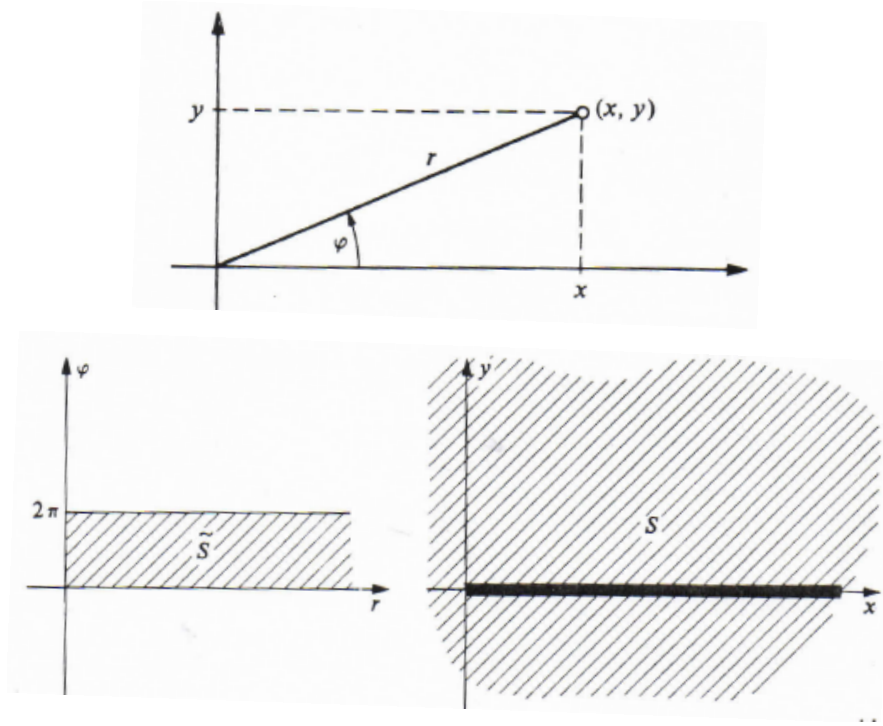


FIGURA 5.8. Ilustración del Ejemplo 5.3.

o sea, $x = g(\tilde{x})$. Puesto que para nuestra aplicación también necesitaremos las derivadas parciales de las funciones $g_1, \dots, g_n$, exigiremos adicionalmente que $\tilde{S}$ sea un conjunto abierto de $\tilde{R} = \mathbb{R}^n$ y $g \in C^1(\tilde{S})$.

Definición 5.3. Sea $S \subset R = \mathbb{R}^n$. Se dice que una función $g : \tilde{R} = \mathbb{R}^n \rightarrow R = \mathbb{R}^n$ define coordenadas nuevas sobre S si se tiene lo siguiente:

1. El conjunto $\tilde{S} = D(g)$ es abierto, y la aplicación $g : \tilde{S} \rightarrow S$ es biyectiva.
2. Se tiene que $g \in C^1(\tilde{S})$, y el determinante de la matriz funcional de g satisface

$$\det \frac{dg}{d\tilde{x}} \neq 0 \quad \text{para todo } \tilde{x} \in \tilde{S}.$$

Ejemplo 5.3 (Coordenadas cilíndricas planas). Consideremos $R = \mathbb{R}^2$. Para un punto $(x, y) \in R$ podemos tratar de describirlo mediante su distancia al origen

$$r = \sqrt{x^2 + y^2}$$

y el ángulo φ formado con el eje x positivo, ver Figura 5.8. Así, poniendo $\tilde{x} = r$ e $\tilde{y} = \varphi$ obtenemos la siguiente aplicación g :

$$g : \begin{cases} x = \tilde{x} \cos \tilde{y} = r \cos \varphi, \\ y = \tilde{x} \sin \tilde{y} = r \sin \varphi. \end{cases}$$

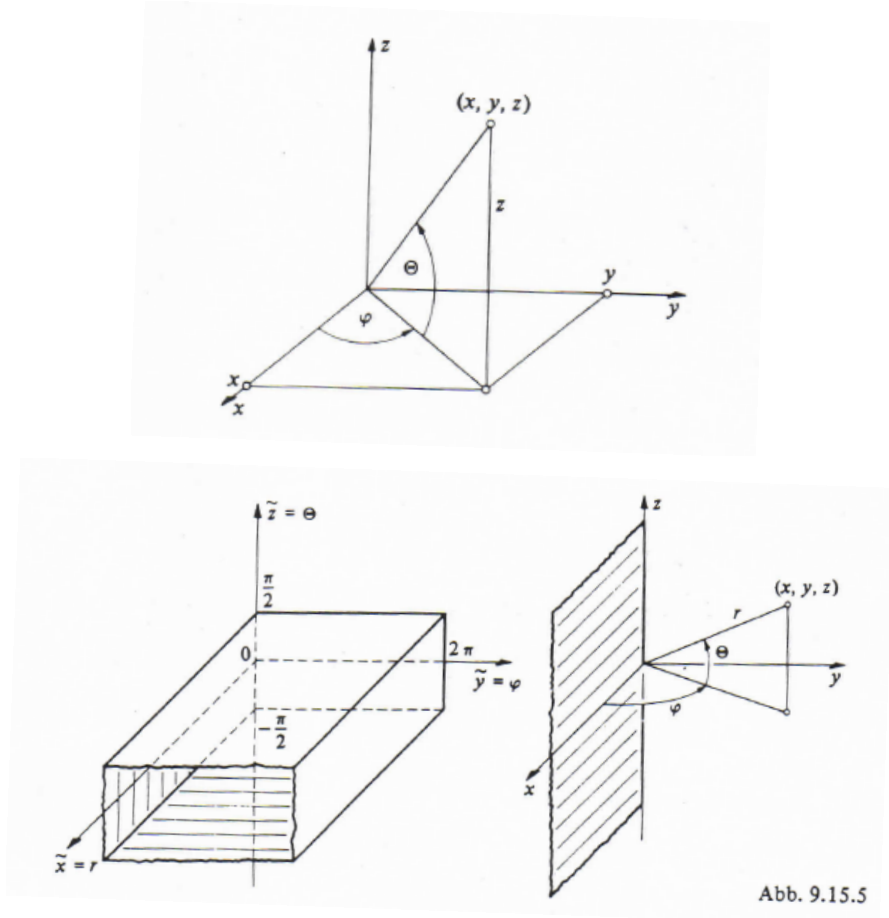


FIGURA 5.9. Ilustración del Ejemplo 5.4.

Sin embargo, tenemos que excluir el origen de $R = \mathbb{R}^2$ dado que allí $\tilde{y} = \tilde{\varphi}$ no es definido unicamente (ver Figura 5.8), además tenemos que considerar la periodicidad de $\cos \varphi$ y $\sin \varphi$ y seleccionar un intervalo donde φ esté bien definido, por ejemplo $0 < \varphi < 2\pi$. Aquí no podemos admitir igualdad puesto que el dominio de g debe ser abierto. En virtud de lo anterior, las coordenadas polares en el plano se definen por

$$g : \begin{cases} x = r \cos \varphi, \\ y = r \sin \varphi, \end{cases} \quad \text{donde } \tilde{S} = D(g) = \{(r, \varphi) \mid 0 < r < \infty, 0 < \varphi < 2\pi\}$$

y $S = R \setminus \{(x, y) \mid x \geq 0, y = 0\}$. Efectivamente sobre S se tiene que

$$\det \frac{dg}{d\tilde{x}} = \begin{vmatrix} \cos \varphi & -r \sin \varphi \\ \sin \varphi & r \cos \varphi \end{vmatrix} = r \neq 0.$$

Ejemplo 5.4 (Coordenadas esféricas espaciales). Sea $R = \mathbb{R}^3$. Para $(x, y, z) \in R$ podemos tratar de describir este punto por su distancia al origen,

$$r = \sqrt{x^2 + y^2 + z^2},$$

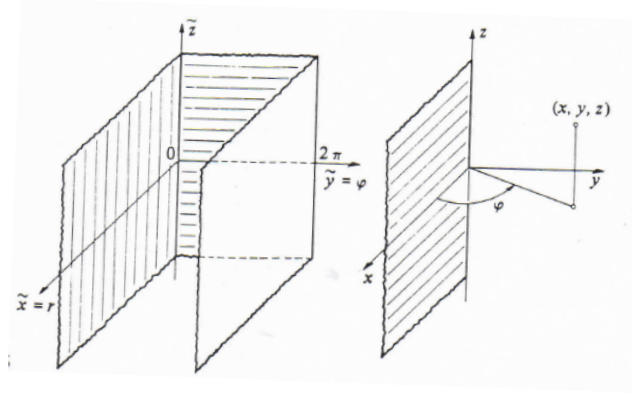


FIGURA 5.10. Ilustración del Ejemplo 5.5.

y dos ángulos con respecto a dos ejes, ver Figura 5.9. Para determinar estos ángulos proyectamos (x, y, z) en la dirección de z al plano (x, y) y luego determinamos los ángulos φ formados con el eje x positivo y θ respecto al plano (x, y) . Entonces, poniendo $\tilde{x} = r$, $\tilde{y} = \varphi$ y $\tilde{z} = \theta$ obtenemos para la aplicación g

$$g : \begin{cases} x = \tilde{x} \cos \tilde{y} \cos \tilde{z} = r \cos \varphi \cos \theta, \\ y = \tilde{x} \sin \tilde{y} \cos \tilde{z} = r \sin \varphi \cos \theta, \\ z = \tilde{x} \sin \tilde{z} = r \sin \theta. \end{cases}$$

Sin embargo, las funciones $\tilde{y} = \varphi$ y $\tilde{z} = \theta$ no son únicamente definidas en el origen, así que hay que excluirlo de $R = \mathbb{R}^3$ (ver Figura 5.9); además tenemos que considerar la periodicidad de las funciones $\cos \varphi$ y $\sin \theta$ y elegir intervalos para los ángulos θ y φ , por ejemplo $0 < \varphi < 2\pi$ y $-\pi/2 < \theta < \pi/2$ (aquí no podemos admitir igualdad dado que el dominio de g debe ser abierto). Entonces, la introducción de coordenadas esféricas puede ser descrita por

$$g : \begin{cases} x = r \cos \varphi \cos \theta, \\ y = r \sin \varphi \cos \theta, \\ z = r \sin \theta, \end{cases}$$

donde

$$\tilde{S} = D(g) = \left\{ (r, \varphi, \theta) \mid 0 < r < \infty, 0 < \varphi < 2\pi, -\frac{\pi}{2} < \theta < \frac{\pi}{2} \right\},$$

y $S = R \setminus \{(x, y, z) \mid x \geq 0, y = 0\}$, donde efectivamente se tiene que

$$\det \frac{dg}{d\tilde{x}} = \begin{vmatrix} \cos \varphi \cos \theta & -r \sin \varphi \cos \theta & -r \cos \varphi \sin \theta \\ \sin \varphi \cos \theta & r \cos \varphi \cos \theta & -r \sin \varphi \sin \theta \\ \sin \theta & 0 & r \cos \theta \end{vmatrix} = r^2 \cos \theta \neq 0.$$

Ejemplo 5.5 (Coordenadas cilíndricas). Sea $R = \mathbb{R}^3$. Se pueden definir las llamadas coordenadas cilíndricas introduciendo coordenadas planas en el plano (x, y) y manteniendo la coordenada z (ver Figura 5.10), es decir, poniendo $\tilde{x} = r$, $\tilde{y} = \varphi$ y $\tilde{z} = z$ obtenemos la

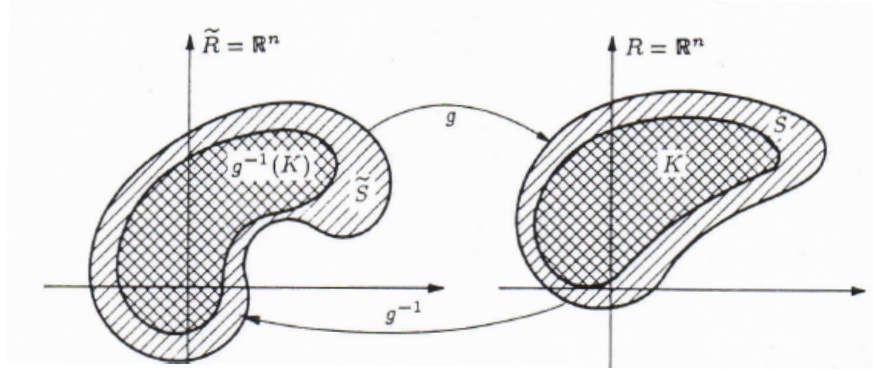


FIGURA 5.11. Ilustración del Teorema 5.3.

aplicación g definida por

$$g : \begin{cases} x = \tilde{x} \cos \tilde{y} = r \cos \varphi, \\ y = \tilde{x} \sin \tilde{y} = r \sin \varphi, \\ z = \tilde{z}, \end{cases}$$

donde

$$\tilde{S} = D(g) = \{(r, \varphi, \tilde{z}) \mid 0 < r < \infty, 0 < \varphi < 2\pi, -\infty < \tilde{z} < \infty\},$$

donde $S = R \setminus \{(x, y, z) \mid x \geq 0, y = 0\}$, donde se tiene que

$$\det \frac{dg}{d\tilde{x}} = \begin{vmatrix} \cos \varphi & -r \sin \varphi & 0 \\ \sin \varphi & r \cos \varphi & 0 \\ 0 & 0 & 1 \end{vmatrix} = r \neq 0.$$

Después de estas consideraciones preliminares formularemos la regla de substitución para integrales n -dimensionales. El problema consiste en la transformación de una integral n -dimensional

$$\int_K f(x) dx$$

en otra integral posiblemente más simple.

Teorema 5.3 (Regla de substitución para integrales n -dimensionales). Sea $\tilde{S} \subset \mathbb{R}^n$ un conjunto abierto, y sea la función $g : \mathbb{R}^n \rightarrow \mathbb{R}^n$ biyectiva sobre $\tilde{S}$ tal que la imagen de $\tilde{S}$ es $S = g(\tilde{S})$, además sea $g \in C^1(\tilde{S})$, y para la matriz funcional (la matriz jacobiana) de g sea

$$\det \frac{dg}{d\tilde{x}}(\tilde{x}) \neq 0 \quad \text{para todo } \tilde{x} \in \tilde{S}.$$

Sea K un subconjunto compacto y medible de S , y sea la función $f : \mathbb{R}^n \rightarrow \mathbb{R}$ continua sobre K . Entonces se tiene que

$$\int_K f(x) dx = \int_{g^{-1}(K)} f(g(\tilde{x})) \left| \det \frac{dg}{d\tilde{x}}(\tilde{x}) \right| d\tilde{x},$$

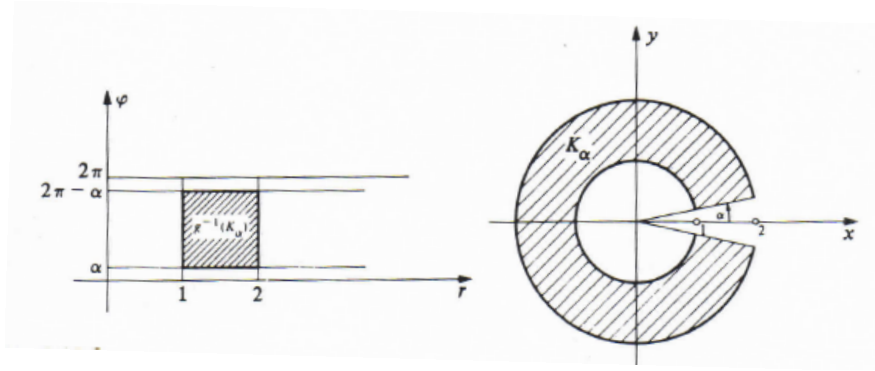


FIGURA 5.12. Ilustración del Ejemplo 5.6.

donde $|\det \frac{dg}{d\tilde{x}}(\tilde{x})|$ (esta cantidad es llamada “el Jacobiano”) denota el valor absoluto de $\det \frac{dg}{d\tilde{x}}(\tilde{x})$.

Este teorema (ver Figura 5.11) no será demostrado aquí, pero ilustraremos su uso a través de algunos ejemplos.

Ejemplo 5.6. Se desea calcular la integral

$$\int_K \sqrt{x^2 + y^2} d(x, y), \quad \text{donde } K = \{(x, y) \mid 1 \leq x^2 + y^2 \leq 4\}$$

(ver Figura 5.12). Aquí se sugiere utilizar las coordenadas polares

$$x = r \cos \varphi,$$

$$y = r \sin \varphi.$$

Ahora, según nuestro teorema, K debería estar contenido en un conjunto sobre el cual hemos definido las coordenadas polares. Pero esto no es así puesto que K contiene puntos del eje x positivo, por lo tanto consideraremos primeramente para un ángulo α el conjunto K_α (ver Figura 5.12).

Ahora podemos aplicar el teorema, dado que K_α es compacto y Riemann-medible. Obtenemos que

$$\begin{aligned} \int_{K_\alpha} \sqrt{x^2 + y^2} d(x, y) &= \int_{g^{-1}(K_\alpha)} r \cdot r d(r, \varphi) \\ &= \int_\alpha^{2\pi-\alpha} \left[\int_1^2 r^2 dr \right] d\varphi \\ &= \int_\alpha^{2\pi-\alpha} \frac{7}{3} d\varphi \\ &= \frac{7}{3} 2(\pi - \alpha). \end{aligned}$$

Ahora tomamos en cuenta que

$$\int_K \sqrt{x^2 + y^2} \, d(x, y) = \int_{K_\alpha} \sqrt{x^2 + y^2} \, d(x, y) + \int_{K \setminus K_\alpha} \sqrt{x^2 + y^2} \, d(x, y).$$

Puesto que

$$\int_{K \setminus K_\alpha} \sqrt{x^2 + y^2} \, d(x, y) \leq 2\mu(K \setminus K_\alpha) \rightarrow 0 \quad \text{cuando } \alpha \rightarrow 0,$$

concluimos que

$$\int_K \sqrt{x^2 + y^2} \, d(x, y) = \frac{14}{3}\pi.$$

Comentamos que las coordenadas planas $x = r \cos \varphi$, $y = r \sin \varphi$ tienen la propiedad importante de que

$$x^2 + y^2 = r^2 \cos^2 \varphi + r^2 \sin^2 \varphi = r^2 \quad \text{para } 0 \leq \varphi \leq 2\pi, \quad r \geq 0.$$

Se recomienda su uso en todas las situaciones si debido a esta relación resultan simplificaciones en la integral que se debe calcular. Esto sucede por ejemplo, si el integrando depende de $x^2 + y^2$, o si el dominio de integración es un círculo o un sector de un círculo.

Específicamente para el Ejemplo 5.6 comentamos que en la práctica no se seguirá cultivando las dudas que nos obligaron a considerar el límite $\alpha \rightarrow 0$. Simplemente se diría en este caso que la imagen recíproca de K es el conjunto

$$g^{-1}(K) = \{(r, \varphi) \mid 1 \leq r \leq 2, 0 \leq \varphi \leq 2\pi\},$$

y se calcularía la integral deseada inmediatamente sobre este rectángulo, es decir como

$$\int_0^{2\pi} \left[\int_1^2 r^2 \, dr \right] d\varphi.$$

Ejemplo 5.7. Queremos calcular el volumen V de

$$K = \{(x, y, z) \mid x \geq 0, y \geq 0, z \geq 0, x^2 + y^2 + z^2 \leq 1\}.$$

Para la computación de la integral

$$V = \iiint_K d(x, y, z)$$

utilizaremos coordenadas polares esféricas:

$$x = r \cos \varphi \cos \theta, \quad y = r \sin \varphi \cos \theta, \quad z = r \sin \theta.$$

Así obtenemos

$$V = \iiint_K d(x, y, z) = \int_0^1 \left\{ \int_0^{\pi/2} \left[\int_0^{\pi/2} r^2 \cos \theta \, d\varphi \right] d\theta \right\} dr$$

$$\begin{aligned}
&= \frac{\pi}{2} \int_0^1 \left[\int_0^{\pi/2} r^2 \cos \theta \, d\theta \right] dr \\
&= \frac{\pi}{2} \int_0^1 r^2 \, dr = \frac{\pi}{6}.
\end{aligned}$$

Comentamos que las coordenadas polares esféricas $x = r \cos \varphi \cos \theta$, $y = r \sin \varphi \cos \theta$, $z = r \sin \theta$ tienen la propiedad importante de que

$$x^2 + y^2 + z^2 = r^2 \quad \text{para } 0 \leq \varphi \leq 2\pi, -\frac{\pi}{2} \leq \theta \leq \frac{\pi}{2} \text{ y } r \geq 0.$$

Se recomienda su uso en todas las situaciones si debido a esta relación resultan simplificaciones en la integral que se debe calcular.

Ejemplo 5.8. *Se desea calcular la integral*

$$\iiint_K x^2 y \, d(x, y, z),$$

donde K es el siguiente segmento de un cilindro:

$$K = \{(x, y, z) \mid x \geq 0, y \geq 0, x^2 + y^2 \leq 1, 0 \leq z \leq 1\}.$$

Utilizando las coordenadas cilíndricas

$$x = r \cos \varphi, \quad y = r \sin \varphi, \quad z = \tilde{z}$$

obtenemos

$$\begin{aligned}
\iiint_K x^2 y \, d(x, y, z) &= \int_0^{\pi/2} \left\{ \int_0^1 \left[\int_0^1 r^4 \cos^2 \varphi \sin \varphi \, d\tilde{z} \right] dr \right\} d\varphi \\
&= \int_0^{\pi/2} \left[\int_0^1 r^4 \cos^2 \varphi \sin \varphi \, dr \right] d\varphi \\
&= \frac{1}{5} \int_0^{\pi/2} \cos^2 \varphi \sin \varphi \, d\varphi \\
&= \left[-\frac{1}{15} \cos^3 \varphi \right]_{\varphi=0}^{\varphi=\pi/2} = \frac{1}{15}.
\end{aligned}$$

Comentamos que las coordenadas cilíndricas $x = r \cos \varphi$, $y = r \sin \varphi$, $z = \tilde{z}$ tienen la propiedad de que

$$x^2 + y^2 = r^2 \quad \text{para } 0 \leq \varphi \leq 2\pi, r \geq 0 \text{ y todo } z = \tilde{z}.$$

Se recomienda su uso en todas las situaciones si debido a esta relación resultan simplificaciones en la integral que se debe calcular.

Ejemplo 5.9. *En la teoría de probabilidades juega un rol importante la integral*

$$\int_0^\infty e^{-x^2} \, dx.$$

Sin embargo, esta integral no puede ser calculada por métodos elementales. Aquí presentaremos un método basado en la integración bidimensional. Para tal efecto consideremos para $R > 0$ los siguientes conjuntos:

$$\begin{aligned} K_R &= \{(x, y) \mid x \geq 0, y \geq 0, x^2 + y^2 \leq R^2\}, \\ Q_R &= \{(x, y) \mid 0 \leq x \leq R, 0 \leq y \leq R\}, \\ K_{R\sqrt{2}} &= \{(x, y) \mid x \geq 0, y \geq 0, x^2 + y^2 \leq 2R^2\}. \end{aligned}$$

Consideremos la función $e^{-(x^2+y^2)}$ sobre estos conjuntos. Utilizando coordenadas polares obtenemos

$$\iint_{K_R} e^{-(x^2+y^2)} d(x, y) = \int_0^{\pi/2} \left[\int_0^R e^{-r^2} r dr \right] d\varphi = \frac{\pi}{4} - \frac{\pi}{4} e^{-R^2}.$$

Análogamente obtenemos

$$\iint_{K_{R\sqrt{2}}} e^{-(x^2+y^2)} d(x, y) = \int_0^{\pi/2} \left[\int_0^{R\sqrt{2}} e^{-r^2} r dr \right] d\varphi = \frac{\pi}{4} - \frac{\pi}{4} e^{-2R^2}.$$

Por otro lado, del Teorema 4.10 se tiene que

$$\begin{aligned} \iint_{Q_R} e^{-(x^2+y^2)} d(x, y) &= \int_0^R \left[\int_0^R e^{-x^2} \cdot e^{-y^2} dx \right] dy \\ &= \left(\int_0^R e^{-x^2} dx \right) \left(\int_0^R e^{-y^2} dy \right) \\ &= \left(\int_0^R e^{-x^2} dx \right)^2, \end{aligned}$$

además, puesto que $K_R \subset Q_R \subset K_{R\sqrt{2}}$,

$$\iint_{K_R} e^{-(x^2+y^2)} d(x, y) \leq \iint_{Q_R} e^{-(x^2+y^2)} d(x, y) \leq \iint_{K_{R\sqrt{2}}} e^{-(x^2+y^2)} d(x, y),$$

por lo tanto

$$\sqrt{\frac{\pi}{4} - \frac{\pi}{4} e^{-R^2}} \leq \int_0^R e^{-x^2} dx \leq \sqrt{\frac{\pi}{4} - \frac{\pi}{4} e^{-2R^2}},$$

lo que significa que

$$\int_0^\infty e^{-x^2} dx = \lim_{R \rightarrow \infty} \int_0^R e^{-x^2} dx = \frac{\sqrt{\pi}}{2}.$$

5.3. Centros de masa y momentos de inercia

Queremos brevemente estudiar algunas aplicaciones del cálculo integral tridimensional en la mecánica. Primeramente consideremos el centro de masa de un conjunto tridimensional.

Definición 5.4. Sean $S \subset \mathbb{R}^3$ un conjunto medible y

$$V = \iiint_S d(x, y, z) \neq 0.$$

En este caso, el punto (x_0, y_0, z_0) con las coordenadas

$$(x_0, y_0, z_0) = \left(\frac{1}{V} \iiint_S x d(x, y, z), \frac{1}{V} \iiint_S y d(x, y, z), \frac{1}{V} \iiint_S z d(x, y, z) \right)$$

se llama centro de masa de S .

Teorema 5.4. Para el centro de masa (x_0, y_0, z_0) de un conjunto S se tiene que

$$\iiint_S (x - x_0) d(x, y, z) = 0, \quad \iiint_S (y - y_0) d(x, y, z) = 0, \quad \iiint_S (z - z_0) d(x, y, z) = 0.$$

Demostración. Según la Definición 5.4 se tiene que

$$\iiint_S (x - x_0) d(x, y, z) = \iiint_S x d(x, y, z) - x_0 \iiint_S d(x, y, z) = 0;$$

se demuestra análogamente que también las demás integrales desaparecen. ■

Ejemplo 5.10. Queremos calcular el centro de masa del conjunto

$$S = \{(x, y, z) \mid x \geq 0, y \geq 0, z \geq 0, x^2 + y^2 + z^2 \leq 1\}$$

ya considerado en el Ejemplo 5.7, donde obtuvimos

$$V = \iiint_S d(x, y, z) = \frac{\pi}{6}.$$

Utilizando coordenadas polares esféricas

$$x = r \cos \varphi \cos \theta, \quad y = r \sin \varphi \cos \theta, \quad z = r \sin \theta$$

obtenemos aquí

$$\begin{aligned} \iiint_S x d(x, y, z) &= \int_0^{\pi/2} \left\{ \int_0^{\pi/2} \left[\int_0^1 r \cos \varphi \cos \theta \cdot r^2 \cos \theta dr \right] d\varphi \right\} d\theta \\ &= \frac{1}{4} \int_0^{\pi/2} \left[\int_0^{\pi/2} \cos^2 \theta \cos \varphi d\varphi \right] d\theta \\ &= \frac{1}{4} \int_0^{\pi/2} \cos^2 \theta d\theta \end{aligned}$$

$$= \frac{1}{8} [\cos \theta \sin \theta + \theta]_{\theta=0}^{\theta=\pi/2} = \frac{\pi}{16}.$$

Mediante cálculos análogos para y_0 y z_0 obtenemos

$$x_0 = \frac{\pi}{16} \cdot \frac{6}{\pi} = \frac{3}{8}, \quad y_0 = \frac{3}{8}, \quad z_0 = \frac{3}{8}.$$

Definición 5.5. Sea $S \subset \mathbb{R}^3$ un conjunto medible.

1. La integral

$$I_x = \iiint_S (y^2 + z^2) \, d(x, y, z)$$

se llama momento de inercia de S con respecto al eje x .

2. La integral

$$I_y = \iiint_S (x^2 + z^2) \, d(x, y, z)$$

se llama momento de inercia de S con respecto al eje y .

3. La integral

$$I_z = \iiint_S (x^2 + y^2) \, d(x, y, z)$$

se llama momento de inercia de S con respecto al eje z .

Ejemplo 5.11. Sea S un cilindro recto y circular del radio R y de la altura h que es simétrico con respecto a la rotación por el eje z . Utilizando coordenadas cilíndricas

$$x = r \cos \varphi, \quad y = r \sin \varphi, \quad z = \tilde{z}$$

obtenemos el siguiente momento de inercia de S con respecto al eje z :

$$I_z = \iiint_S (x^2 + y^2) \, d(x, y, z) = \int_0^R \left\{ \int_0^{2\pi} \left[\int_{-h/2}^{h/2} r^3 \, d\tilde{z} \right] d\varphi \right\} dr = \frac{\pi R^4 h}{2}.$$

Ejemplo 5.12. Se considera el conjunto

$$S = \left\{ (x, y, z) \mid -\frac{a}{2} \leq x \leq \frac{a}{2}, -\frac{b}{2} \leq y \leq \frac{b}{2}, -\frac{c}{2} \leq z \leq \frac{c}{2} \right\}.$$

Queremos calcular el momento inercial con respecto al eje x . Aquí obtenemos

$$I_x = \iiint_S (y^2 + z^2) \, d(x, y, z) = \int_{-c/2}^{c/2} \left\{ \int_{-b/2}^{b/2} \left[\int_{-a/2}^{a/2} (y^2 + z^2) \, dx \right] dy \right\} dz = \frac{abc}{12} (b^2 + c^2).$$

Análisis vectorial y teoremas integrales

6.1. Curvas en $\mathbb{R}^n$ y el vector tangencial

Definición 6.1. Sea $f : \mathbb{R} \rightarrow \mathbb{R}^n$, y sea f continua sobre $[a, b]$.

a) El conjunto

$$K = \{x \in \mathbb{R}^n \mid x = f(t) = (f_1(t), \dots, f_n(t)), t \in [a, b]\}$$

se llama curva en $\mathbb{R}^n$; $(f, [a, b])$ es una parametrización de K ; y si $f(a) = f(b)$, entonces la curva K se llama cerrada.

b) Si se tiene que $f(t_2) = f(t_1)$ para $a \leq t_1 \leq t_2 \leq b$ solamente para $t_1 = t_2$ o $t_1 = a$ y $t_2 = b$, entonces K se llama curva de Jordan.

Evidentemente, una curva de Jordan es una curva que no posee puntos dobles, con la posible excepción de $f(a)$ y $f(b)$.

Si f es una función real y continua definida sobre $[a, b]$, entonces el conjunto de puntos al cual hasta ahora nos hemos referido como grafo o diagrama de f es una curva en el sentido de la Definición 6.1. Solamente hay que elegir $x = f_1(t) = t$ y $y = f_2(t) = f(t)$, ver Figura 6.2.

Por otro lado, una parametrización induce sobre K una dirección, dado que si t asume los valores del intervalo $[a, b]$ desde a hasta b , los puntos $x = f(t)$ asumen las posiciones desde $f(a)$ hasta $f(b)$, por lo tanto $f(a)$ y $f(b)$ son los puntos inicial y terminal de K .

Muy frecuentemente representaremos una curva vectorialmente, es decir K es el conjunto de los puntos terminales de los vectores

$$\vec{f}(t) = \{f_1(t), \dots, f_n(t)\},$$

ver Figura 6.3.

Ejemplo 6.1. Sean $f_1(t) = \cos t$ y $f_2(t) = \sin t$ para $t \in [0, \pi]$. En este caso K es una curva de Jordan; específicamente, K es la mitad superior del círculo unitario.

Ejemplo 6.2. Sean $f_1(t) = t$ y $f_2(t) = \sqrt{1 - t^2}$ para $t \in [-1, 1]$. Entonces K es una curva de Jordan; específicamente, K nuevamente es la mitad superior del círculo unitario.

Ejemplo 6.3. Para $r > 0$ y $c > 0$ sean $f_1(t) = r \cos t$, $f_2(t) = r \sin t$ y $f_3(t) = ct$ para $t \in [a, b]$. La curva K se llama espiral o curva helicoidal cilíndrica, ver Figura 6.4.

Los Ejemplos 6.1 y 6.2 ilustran que una curva K puede tener varias parametrizaciones. Para algunos propósitos es necesario estudiar parametrizaciones equivalentes de una curva.

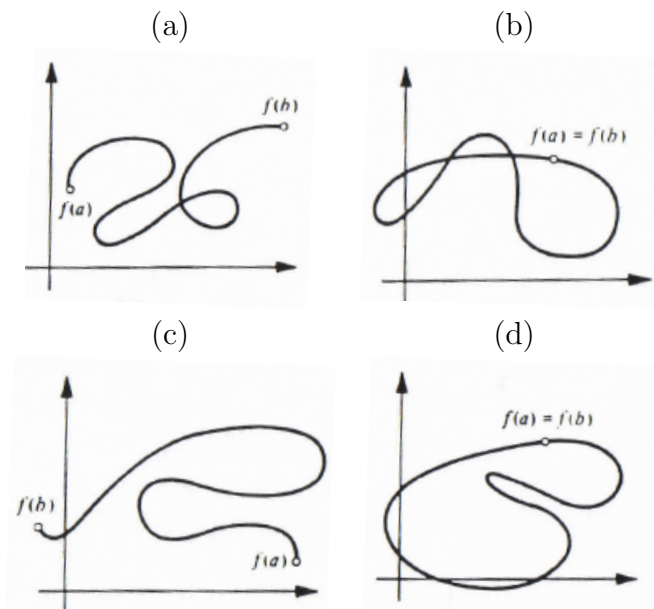


FIGURA 6.1. Ilustración de la Definición 6.1: (a) curva, (b) curva cerrada, (c) curva de Jordan, (d) curva de Jordan cerrada.

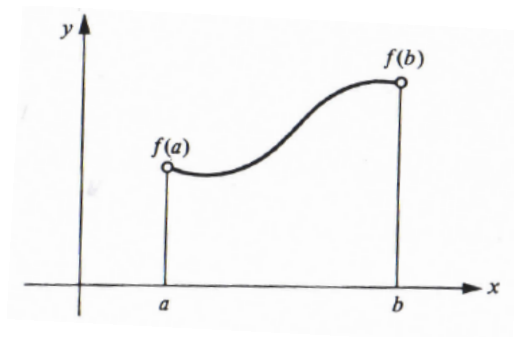


FIGURA 6.2. Curva definida mediante una función $f : [a, b] \rightarrow \mathbb{R}$.

Definición 6.2. Sean $(f, [a, b])$ y $(g, [c, d])$ parametrizaciones de una curva K . Si existe una función $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ tal que

$$\varphi \in C^1[a, b], \quad \varphi'(t) > 0 \text{ sobre } [a, b], \quad (6.1)$$

$$\varphi(a) = c, \quad \varphi(b) = d, \quad (6.2)$$

$$f(t) = g(\varphi(t)) \quad \text{para todo } t \in [a, b], \quad (6.3)$$

entonces las parametrizaciones $(f, [a, b])$ y $(g, [c, d])$ se dicen equivalentes, ver Figura 6.5.

Un concepto muy importante es el del vector tangencial de una curva. Para motivar este concepto, consideremos una curva K con la parametrización $(f, [a, b])$. Para puntos $t \in [a, b]$

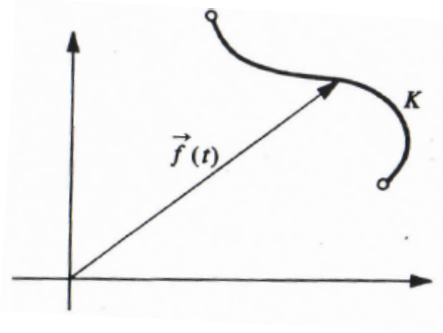


FIGURA 6.3. Representación vectorial de una curva.

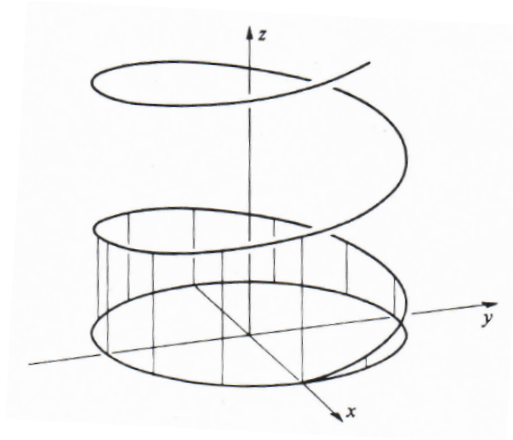


FIGURA 6.4. Espiral.

y $t + h \in [a, b]$ con $h \neq 0$ formamos el vector de diferencia

$$\vec{f}(t+h) - \vec{f}(t) = \{f_1(t+h) - f_1(t), \dots, f_n(t+h) - f_n(t)\},$$

ver Figura 6.6. Dividiendo por h obtenemos el vector

$$\frac{\vec{f}(t+h) - \vec{f}(t)}{h} = \left\{ \frac{f_1(t+h) - f_1(t)}{h}, \dots, \frac{f_n(t+h) - f_n(t)}{h} \right\}.$$

Ahora, si cada una de las funciones $f_1, \dots, f_n$ es diferenciable en la posición t , esta expresión converge cuando $h \rightarrow 0$ al vector límite

$$\vec{f}'(t) = \frac{d\vec{f}}{dt}(t) = \{f'_1(t), \dots, f'_n(t)\}.$$

La longitud de este vector es

$$\|\vec{f}'(t)\| = \sqrt{\sum_{\nu=1}^n (f'_\nu(t))^2},$$

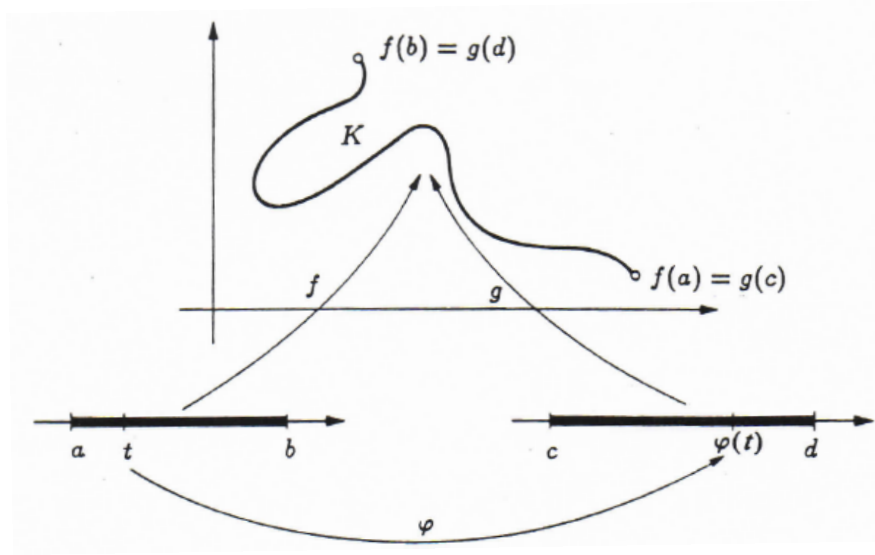


FIGURA 6.5. Parametrizaciones equivalentes $(f, [a, b])$ y $(g, [c, d])$ de una curva K .

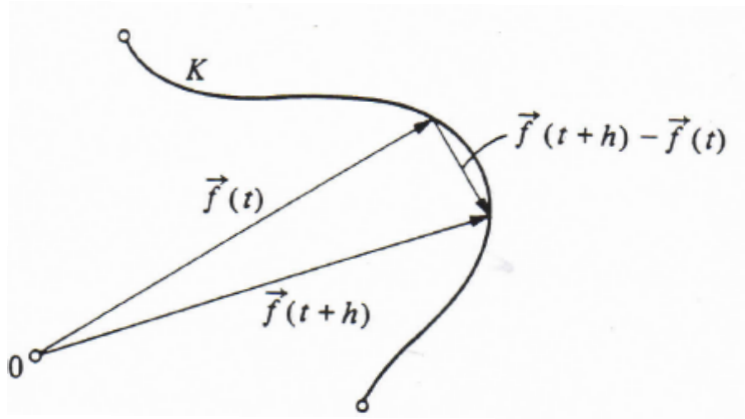


FIGURA 6.6. El vector de diferencia $\vec{f}'(t+h) - \vec{f}'(t)$.

y evidentemente $\vec{f}'(t) = 0$ si y sólo si simultaneamente $f'_1(t) = 0, \dots, f'_n(t) = 0$. Si $\vec{f}'(t) \neq 0$, entonces el vector $\vec{f}'(t)/\|\vec{f}'(t)\|$ define una dirección en $\mathbb{R}^n$.

Definición 6.3. Sea K una curva de $\mathbb{R}^n$ con la parametrización $(f, [a, b])$, donde $f = (f_1, \dots, f_n) : \mathbb{R} \rightarrow \mathbb{R}^n$.

1. Si el vector $\vec{f}'(t) = \{f'_1(t), \dots, f'_n(t)\}$ existe para un $t \in [a, b]$ y se tiene que $\vec{f}'(t) \neq \vec{0}$, entonces

$$\vec{T}_f(t) = \frac{1}{\|\vec{f}'(t)\|} \vec{f}'(t)$$

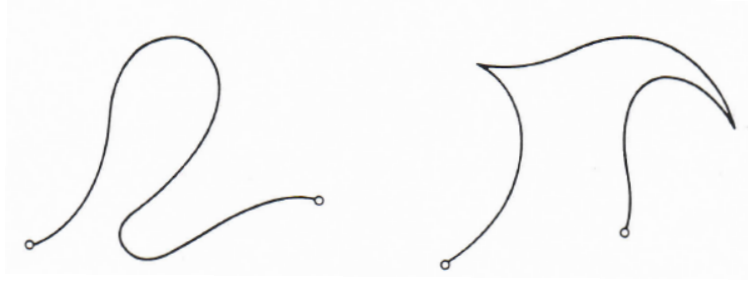


FIGURA 6.7. Una curva suave (izquierda) y una curva suave por trozos (derecha).

se llama vector tangencial de K en el punto $f(t)$ (con respecto a la parametrización $(f, [a, b])$).

2. La recta

$$\{x \mid x = f(t) + \lambda \vec{f}'(t), \lambda \in \mathbb{R}\}$$

se llama tangente de K en el punto $f(t)$.

Sean $(f, [a, b])$ y $(g, [c, d])$ parametrizaciones equivalentes de la curva K , y sean $f = (f_1, \dots, f_n)$ y $g = (g_1, \dots, g_n)$. Puesto que

$$f'_i(t) = g'_i(\varphi(t)) \cdot \varphi'(t)$$

se tiene que

$$\vec{T}_f(t) = \frac{\vec{f}'(t)}{\|\vec{f}'(t)\|} = \frac{g'(\varphi(t)) \cdot \varphi'(t)}{\|g'(\varphi(t)) \cdot \varphi'(t)\|} = \frac{\vec{g}'(\varphi(t))}{\|\vec{g}'(\varphi(t))\|} = \vec{T}_g(\varphi(t))$$

siempre que el vector tangencial $\vec{T}_f(t)$ existe, es decir, $\vec{f}'(t) \neq \vec{0}$. Esto significa que el vector tangencial no cambia si pasamos a una parametrización equivalente.

El concepto del vector tangencial nos permite definir curvas suaves y curvas suaves por trozos, ver Figura 6.7.

Definición 6.4. Sea K una curva en $\mathbb{R}^n$ con la parametrización $(f, [a, b])$, y sea $f = (f_1, \dots, f_n) : \mathbb{R} \rightarrow \mathbb{R}^n$.

1. La curva K se llama suave si $\vec{T}_f(t)$ existe sobre $[a, b]$ y las funciones f'_ν , $\nu = 1, \dots, n$, son continuas sobre $[a, b]$.
2. La curva K se llama suave por trozos si existe una partición $P = \{t_0, \dots, t_m\}$ de $[a, b]$ tal que K es suave sobre cada intervalo $[t_{k-1}, t_k]$, $k = 1, \dots, m$.

6.2. Funciones de variación acotada

Para la discusión de curvas en $\mathbb{R}^n$ necesitamos el concepto de las funciones de variación acotada, el cual también es importante en otros áreas del análisis.

Definición 6.5. Sea la función $f : \mathbb{R} \rightarrow \mathbb{R}$ definida sobre $[a, b]$.

1. Si existe una constante $M > 0$ tal que para todas las particiones $P = \{x_0, x_1, \dots, x_n\}$ de $[a, b]$ se tiene que

$$V_P(f) := \sum_{i=1}^n |f(x_i) - f(x_{i-1})| \leq M,$$

entonces la función f se llama de variación acotada sobre $[a, b]$. Escribimos también $f \in BV[a, b]$.

2. Para $f \in BV[a, b]$ definimos la cantidad

$$\overset{b}{V}_a(f) = \sup_P V_P(f),$$

llamada la variación total de f sobre $[a, b]$. Formalmente, se define

$$\overset{a}{V}_a(f) = 0.$$

Evidentemente, se tiene que

$$\overset{b}{V}_a(f) = \overset{b}{V}_a(-f).$$

Ejemplo 6.4. La función $f(x) = x$ es de variación acotada sobre cada intervalo $[a, b]$. Para ver eso, sea $P = \{x_0, x_1, \dots, x_n\}$ una partición arbitraria de $[a, b]$. Entonces

$$V_P(f) = \sum_{i=1}^n |x_i - x_{i-1}| = \sum_{i=1}^n (x_i - x_{i-1}) = x_n - x_0 = b - a.$$

Ejemplo 6.5. Consideremos la siguiente función, que es continua sobre $\mathbb{R}$:

$$f(x) = \begin{cases} x \sin(\pi/x) & \text{para } x \neq 0, \\ 0 & \text{para } x = 0. \end{cases}$$

Esta función no es de variación acotada sobre el intervalo $[0, 2]$. Para ver eso, consideremos para un $n \in \mathbb{N}$ la partición $P = \{x_0, \dots, x_n\}$ de $[0, 2]$ dada por

$$x_0 = 0; \quad x_i = \frac{2}{2n+1-2i}, \quad i = 1, \dots, n.$$

En este caso, se tiene lo siguiente para $n > 2$:

$$\begin{aligned} V_P(f) &= \sum_{i=1}^n |f(x_i) - f(x_{i-1})| \\ &\geq \sum_{i=2}^n |f(x_i) - f(x_{i-1})| \\ &= \sum_{i=2}^n \left| \frac{2}{2n+1-2i} \sin\left(\frac{2n+1-2i}{2}\pi\right) - \frac{2}{2n+3-2i} \sin\left(\frac{2n+3-2i}{2}\pi\right) \right| \\ &= \sum_{i=2}^n \left| \sin\left((n-i)\pi + \frac{\pi}{2}\right) \right| \left(\frac{2}{2n+1-2i} + \frac{2}{2n+3-2i} \right) \end{aligned}$$

$$> \sum_{i=2}^n \frac{1}{2n+1-2i} = \sum_{k=0}^{n-2} \frac{1}{1+2k}.$$

Pero en virtud de la divergencia de la serie armónica

$$\sum_{k=0}^{\infty} \frac{1}{1+2k},$$

esta última expresión puede ser arbitrariamente grande, por lo tanto $f \notin BV[0, 2]$. Este ejemplo demuestra que una función continua no necesariamente debe ser de variación acotada. Vice versa, una función de variación acotada no necesariamente debe ser continua, como ilustra el siguiente ejemplo.

Ejemplo 6.6. Consideremos sobre $[-1, 1]$ la función discontinua

$$f(x) = \begin{cases} 0 & \text{si } x \in [-1, 0], \\ 1 & \text{si } x \in (0, 1]. \end{cases}$$

Evidentemente, para cada partición P de $[-1, 1]$ se tiene que $V_P(f) = 1$, entonces f es de variación acotada.

Teorema 6.1. Sea la función f definida sobre $[a, b]$.

1. Si f es monótona sobre $[a, b]$, entonces $f \in BV[a, b]$.
2. Si f es continua sobre $[a, b]$, diferenciable sobre (a, b) y si f' es acotada, entonces $f \in BV[a, b]$.

Demostración. Sea $P = \{x_0, \dots, x_n\}$ una partición arbitraria de $[a, b]$.

1. Sea f monótona sobre $[a, b]$. Sin pérdida de la generalidad suponemos que f es creciente (si f es decreciente, la demostración es análoga). La primera afirmación es una consecuencia de

$$V_P(f) = \sum_{i=1}^n |f(x_i) - f(x_{i-1})| = \sum_{i=1}^n (f(x_i) - f(x_{i-1})) = f(b) - f(a).$$

2. Sea f continua sobre $[a, b]$, diferenciable sobre (a, b) y con una constante K sea $|f'(x)| \leq K$ para $x \in (a, b)$. Según el Teorema del Valor Intermedio del Cálculo Diferencial existen puntos $\xi_i \in (x_{i-1}, x_i)$, $i = 1, \dots, n$, tales que

$$\begin{aligned} V_P(f) &= \sum_{i=1}^n |f(x_i) - f(x_{i-1})| = \sum_{i=1}^n \left| \frac{f(x_i) - f(x_{i-1})}{x_i - x_{i-1}} \right| (x_i - x_{i-1}) \\ &= \sum_{i=1}^n |f'(\xi_i)| (x_i - x_{i-1}) \leq K \sum_{i=1}^n (x_i - x_{i-1}) = K(b - a). \end{aligned}$$

■

El Ejemplo 6.5 demuestra que una función acotada no necesariamente debe ser de variación acotada. Sin embargo, una función de variación acotada siempre es acotada.

Teorema 6.2. *Sea la función f de variación acotada sobre $[a, b]$. Entonces f es acotada sobre $[a, b]$.*

Demostración. Para un $x \in [a, b]$ arbitrario se tiene que

$$|f(b) - f(x)| + |f(x) - f(a)| \leq \overset{b}{V}_a(f),$$

lo que implica que

$$\begin{aligned} |f(x)| &\leq |f(x) - f(a)| + |f(a)| \\ &\leq |f(b) - f(x)| + |f(x) - f(a)| + |f(a)| \\ &\leq \overset{b}{V}_a(f) + |f(a)|, \end{aligned}$$

por lo tanto f es acotada sobre $[a, b]$. ■

Teorema 6.3. *Sean $f \in BV[a, b]$ y $g \in BV[a, b]$. Entonces (a) $f + g \in BV[a, b]$, (b) $f - g \in BV[a, b]$, (c) $f \cdot g \in BV[a, b]$ y (d) $f/g \in BV[a, b]$ si $g(x) \geq \delta > 0$.*

Demostración. Tarea. ■

Teorema 6.4. *Sean $f \in BV[a, b]$ y $c \in [a, b]$. Entonces se tiene que*

1. $f \in BV[a, c]$ y $f \in BV[c, b]$,
2. $\overset{b}{V}_a(f) = \overset{c}{V}_a(f) + \overset{b}{V}_c(f)$.

Demostración. Si $c = a$ o $c = b$, no hay nada que demostrar; sea entonces $a < c < b$.

1. Sean P_a^c y P_c^b particiones de $[a, c]$ y $[c, b]$, respectivamente. Entonces $P_a^b = P_a^c \cup P_c^b$ es una partición de $[a, b]$, y sabemos que

$$V_{P_a^c}(f) + V_{P_c^b}(f) = V_{P_a^b}(f) \leq \overset{b}{V}_a(f).$$

En particular, esto implica que

$$V_{P_a^c}(f) \leq \overset{b}{V}_a(f), \quad V_{P_c^b}(f) \leq \overset{b}{V}_c(f),$$

es decir $f \in BV[a, c]$ y $f \in BV[c, b]$; además se tiene que

$$\overset{c}{V}_a(f) + \overset{b}{V}_c(f) \leq \overset{b}{V}_a(f).$$

2. Para la demostración de la desigualdad opuesta consideremos una partición $P = \{x_0, \dots, x_n\}$ de $[a, b]$. La partición $P_a^b = P \cup \{c\}$ nuevamente es una partición de $[a, b]$ y genera particiones P_a^c y P_c^b de $[a, c]$ y $[c, b]$, respectivamente. Para un índice k apropiado se tiene que

$$x_{k-1} \leq c < x_k$$

y

$$|f(x_k) - f(x_{k-1})| \leq |f(x_k) - f(c)| + |f(c) - f(x_{k-1})|.$$

Insertando esto en $V_P(f)$ obtenemos

$$V_P(f) \leq V_{P_a^b}(f) = V_{P_a^c}(f) + V_{P_c^b}(f) \leq \bar{V}_a^c(f) + \bar{V}_c^b(f).$$

Dado que P es una partición arbitraria, obtenemos de esto

$$\bar{V}_a^b(f) \leq \bar{V}_a^c(f) + \bar{V}_c^b(f).$$

■

Teorema 6.5. Sea $f \in BV[a, b]$. Entonces para la función

$$V(x) := \bar{V}_a^x(f), \quad x \in [a, b]$$

sabemos que

1. V es monótona creciente sobre $[a, b]$,
2. $V + f$ es monótona creciente sobre $[a, b]$,
3. $V - f$ monótona creciente sobre $[a, b]$.

Demostración. Sea $a < x_1 < x_2 < b$.

1. De acuerdo al Teorema 6.4,

$$\bar{V}_a^{x_2}(f) = \bar{V}_a^{x_1}(f) + \bar{V}_{x_1}^{x_2}(f),$$

es decir

$$V(x_2) - V(x_1) = \bar{V}_a^{x_2}(f) - \bar{V}_a^{x_1}(f) = \bar{V}_{x_1}^{x_2}(f) \geq 0.$$

2. Para la partición trivial $\{x_1, x_2\}$ de $[x_1, x_2]$ sabemos que

$$f(x_2) - f(x_1) \leq |f(x_2) - f(x_1)| \leq \bar{V}_{x_1}^{x_2}(f),$$

por lo tanto

$$(V(x_2) - f(x_2)) - (V(x_1) - f(x_1)) = \bar{V}_{x_1}^{x_2}(f) - (f(x_2) - f(x_1)) \geq 0.$$

3. La demostración que $V + f$ es monótona creciente es análoga.

■

Ahora podemos ofrecer la siguiente caracterización importante de las funciones de variación acotada.

Teorema 6.6. La función f es de variación acotada sobre $[a, b]$ si y sólo si posee una representación $f(x) = f_1(x) - f_2(x)$, donde ambas funciones f_1 y f_2 son monótonas crecientes sobre $[a, b]$.

Demostración.

1. Si $f(x) = f_1(x) - f_2(x)$, donde ambas funciones f_1 y f_2 son monótonas crecientes sobre $[a, b]$, según los Teoremas 6.1 y 6.3 también f es de variación acotada.

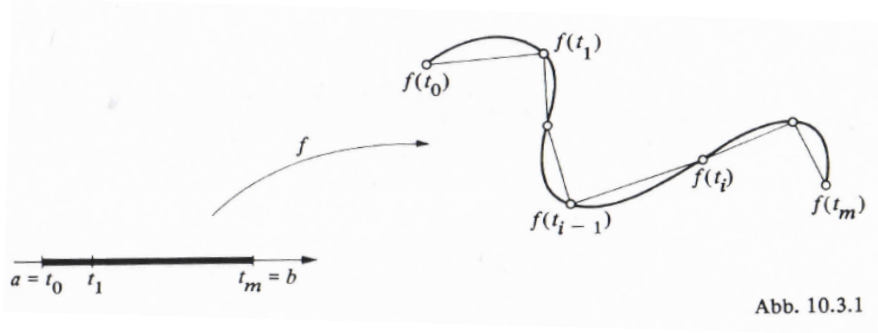


FIGURA 6.8. Trazado poligonal por los puntos $f(t_0), f(t_1), \dots, f(t_m)$ de la curva K .

2. Por otro lado, sea f de variación acotada, entonces las funciones f_1 y f_2 definidas por

$$f_1(x) = \frac{V(x) + f(x)}{2}, \quad f_2(x) = \frac{V(x) - f(x)}{2}$$

son ambas monótonas crecientes (según el Teorema 6.5), y $f(x) = f_1(x) - f_2(x)$. ■

El Teorema 6.6 implica, en particular, que una función $f \in BV[a, b]$ es Riemann-integrable sobre $[a, b]$.

6.3. La longitud de una curva

Para el caso especial de una curva K que es el grafo de una función $f : \mathbb{R} \rightarrow \mathbb{R}$ ya definimos el concepto de la longitud de arco. Ahora vamos a generalizar este concepto a curvas arbitrarias de $\mathbb{R}^n$. Para tal efecto, consideremos una curva $K \subset \mathbb{R}^n$ con la parametrización $(f, [a, b])$. Para una partición arbitraria $P = \{t_0, \dots, t_m\}$ de $[a, b]$ consideremos el trazado poligonal por los puntos $f(t_0), f(t_1), \dots, f(t_m)$, ver Figura 6.8. La longitud de este trazado poligonal es

$$L_P(K) = \sum_{i=1}^m d(f(t_{i-1}), f(t_i)).$$

Ahora si dejamos tender $\|P\| \rightarrow 0$, entonces nuestra expectativa es que para una curva K “razonable”, los trazados poligonales aproximan la curva K y que las longitudes tienden hacia un límite, el cual vamos a considerar como “longitud de K ”. Esto se precisa en la siguiente definición.

Definición 6.6. Sea K una curva en $\mathbb{R}_+^n$ con la parametrización $(f, [a, b])$.

1. Si existe una constante M tal que para toda partición $P = \{t_0, \dots, t_m\}$ de $[a, b]$ se tiene que

$$L_P(K) = \sum_{i=1}^m d(f(t_{i-1}), f(t_i)) \leq M, \quad (6.4)$$

entonces la curva K se llama rectificable.

2. Si K es rectificable, entonces

$$L(K) = \sup_P L_P(K)$$

se llama la longitud de arco de K (con respecto a la parametrización $(f, [a, b])$).

Si la función $f : \mathbb{R} \rightarrow \mathbb{R}^n$ está dada por $f = (f_1, \dots, f_n)$, entonces la relación (6.4) puede ser escrita como

$$L_P(K) = \sum_{i=1}^m d(f(t_{i-1}), f(t_i)) = \sum_{i=1}^m \sqrt{\sum_{k=1}^n (f_k(t_i) - f_k(t_{i-1}))^2} \leq M.$$

Comentamos que en el caso especial $K = C_f$ en $\mathbb{R}^2$, es decir, cuando la curva está dada a través de una función $y = f(x)$, las definiciones de la rectificabilidad y de la longitud de arco coinciden con los conceptos correspondientes del cálculo de funciones de una variable.

Teorema 6.7. Sea K una curva en $\mathbb{R}^n$ con la parametrización $(f, [a, b])$, $f = (f_1, \dots, f_n) : \mathbb{R} \rightarrow \mathbb{R}^n$. Entonces la curva K es rectificable si y sólo si cada función de coordenada f_k , $k = 1, \dots, n$, es de variación acotada sobre $[a, b]$.

Demostración.

1. Sea $f_k \in BV[a, b]$ para $k = 1, \dots, n$ y $P = \{t_0, \dots, t_m\}$ una partición arbitraria de $[a, b]$. Entonces

$$d(f(t_{i-1}), f(t_i)) = \sqrt{\sum_{k=1}^n (f_k(t_i) - f_k(t_{i-1}))^2} \leq \sum_{k=1}^n |f_k(t_i) - f_k(t_{i-1})|$$

para $i = 1, \dots, m$, luego

$$\begin{aligned} L_P(K) &= \sum_{i=1}^m d(f(t_{i-1}), f(t_i)) \leq \sum_{i=1}^m \left(\sum_{k=1}^n |f_k(t_i) - f_k(t_{i-1})| \right) \\ &= \sum_{k=1}^n \left(\sum_{i=1}^m |f_k(t_i) - f_k(t_{i-1})| \right) \leq \sum_{k=1}^n V_a^b(f_k), \end{aligned}$$

lo que implica que K es rectificable.

2. Sea K rectificable y $P = \{t_0, \dots, t_m\}$ una partición arbitraria de $[a, b]$. Entonces, para $k = 1, \dots, n$ e $i = 1, \dots, m$ se tiene que

$$|f_k(t_i) - f_k(t_{i-1})| \leq d(f(t_{i-1}), f(t_i)).$$

Según la Definición 6.6, esto significa que

$$\sum_{i=1}^m |f_k(t_i) - f_k(t_{i-1})| \leq \sum_{i=1}^m d(f(t_{i-1}), f(t_i)) \leq L(K),$$

por lo tanto, $f_k \in BV[a, b]$ para $k = 1, \dots, n$. ■

Teorema 6.8. Sea $K \subset \mathbb{R}^n$ una curva con la parametrización $(f, [a, b])$, y sea $c \in (a, b)$. Entonces K es rectificable si y sólo si cada una de las curvas parciales K_1 con la parametrización $(f, [a, c])$ y K_2 con la parametrización $(f, [c, b])$ es rectificable, además se tiene que

$$L(K) = L(K_1) + L(K_2).$$

Demostración.

1. Sea la curva K rectificable.

a) Sea P_1 una partición de $[a, c]$ y P_2 una partición de $[c, b]$. Entonces $P = P_1 \cup P_2$ es una partición de $[a, b]$, y sabemos que

$$L_{P_1}(K_1) + L_{P_2}(K_2) = L_P(K) \leq L(K).$$

Esto significa $L_{P_1}(K_1) \leq L(K)$ y $L_{P_2}(K_2) \leq L(K)$, es decir ambas curvas K_1 y K_2 son rectificables, además se debe tener que

$$L(K_1) + L(K_2) \leq L(K).$$

b) Sea $P = \{t_0, \dots, t_m\}$ una partición de $[a, b]$. La partición $P \cup \{c\}$ entrega particiones P_1 de $[a, c]$ y P_2 de $[c, b]$. Entonces para un índice i apropiado, $1 \leq i \leq m$, se tiene que $t_{i-1} \leq c < t_i$, luego

$$d(f(t_{i-1}), f(t_i)) \leq d(f(t_{i-1}), f(c)) + d(f(t_i), f(c)).$$

Insertando eso en $L_P(K)$ obtenemos

$$L_P(K) \leq L_{P_1}(K_1) + L_{P_2}(K_2) \leq L(K_1) + L(K_2)$$

y de eso, a su vez,

$$L(K) = \sup_P L_P(K) \leq L(K_1) + L(K_2).$$

De (a) y (b) sigue la rectificabilidad de K_1 y K_2 tanto que

$$L(K) = L(K_1) + L(K_2).$$

2. Si K_1 y K_2 son ambas rectificables, podemos de manera similar demostrar la rectificabilidad de K . La validez de $L(K) = L(K_1) + L(K_2)$ es una consecuencia de (1). ■

El siguiente teorema presenta una fórmula que permite la computación explícita de la longitud de arco de una curva en un caso particular.

Teorema 6.9. Sea $K \subset \mathbb{R}^n$ una curva suave por trozos con la parametrización $(f, [a, b])$. Sea $f = (f_1, \dots, f_n)$. Entonces K es rectificable, y para su longitud de arco tenemos que

$$L(K) = \int_a^b \sqrt{\sum_{k=1}^n (f'_k(t))^2} dt.$$

Demostración. Según el Teorema 6.8 es suficiente demostrar el Teorema 6.9 para el caso de una curva K suave. Esto significa que según la Definición 6.4, las funciones f'_k son continuas sobre $[a, b]$ para $k = 1, \dots, n$, por lo tanto podemos aplicar el Teorema 6.7 para concluir que $f_k \in BV[a, b]$ para $k = 1, \dots, n$, y de acuerdo al Teorema 6.7 sigue la rectificabilidad de K .

1. Sea $\varepsilon > 0$ dado.

a) La función

$$\varphi(t) = \sqrt{\sum_{k=1}^n (f'_k(t))^2}, \quad t \in [a, b]$$

es continua y por lo tanto integrable sobre $[a, b]$. Sea $P = \{t_0, t_1, \dots, t_m\}$ una partición de $[a, b]$ primeramente arbitraria, entonces podemos formar las sumas de Riemann

$$S_P(\varphi, \tau) = \sum_{i=1}^m \varphi(t_i)(t_i - t_{i-1}) = \sum_{i=1}^m \sqrt{\sum_{k=1}^n (f'_k(t_i))^2} \cdot (t_i - t_{i-1}).$$

Ahora podemos elegir $\delta_0 > 0$ de tal modo que para todas estas particiones con $\|P\| < \delta_0$ se tiene que

$$\left| \int_a^b \varphi(t) dt - S_P(\varphi, \tau) \right| < \frac{\varepsilon}{2}. \quad (6.5)$$

b) Para $k = 1, \dots, n$ las funciones f'_k son uniformemente continuas sobre $[a, b]$, por lo tanto podemos elegir un δ con $0 < \delta < \delta_0$ de la manera que para todo $k = 1, \dots, n$ y todo $t, \tilde{t} \in [a, b]$ con $|t - \tilde{t}| < \delta$ se tiene que

$$|f'_k(t) - f'_k(\tilde{t})| < \frac{\varepsilon}{2\sqrt{n}(b-a)}.$$

2. Sea ahora la partición P elegida tal que $\|P\| < \delta$. Para $i = 1, \dots, m$ sabemos que

$$\begin{aligned} (d(f(t_{i-1}), f(t_i)))^2 &= \sum_{k=1}^n (f_k(t_i) - f_k(t_{i-1}))^2 \\ &= \sum_{k=1}^n \left(\frac{f_k(t_i) - f_k(t_{i-1})}{t_i - t_{i-1}} \right)^2 (t_i - t_{i-1})^2. \end{aligned}$$

Según el Teorema del Valor Intermedio del Cálculo Diferencial existe para cada $i = 1, \dots, m$ y cada $k = 1, \dots, n$ un $\tau_k^i \in (t_{i-1}, t_i)$ tal que

$$\frac{f_k(t_i) - f_k(t_{i-1})}{t_i - t_{i-1}} = f'_k(\tau_k^i),$$

y se tiene que

$$L_P(K) = \sum_{i=1}^m d(f(t_{i-1}), f(t_i)) = \sum_{i=1}^m \sqrt{\sum_{k=1}^n (f'_k(\tau_k^i))^2} \cdot (t_i - t_{i-1}).$$

Utilizando la desigualdad de Minkowski, obtenemos ahora

$$|S_P(\varphi, \tau) - L_P(K)|$$

$$\begin{aligned}
&= \left| \sum_{i=1}^m \left(\sqrt{\sum_{k=1}^n (f'_k(t_i))^2} - \sqrt{\sum_{k=1}^n (f'_k(\tau_k^i))^2} \right) (t_i - t_{i-1}) \right| \\
&\leq \sum_{i=1}^m \sqrt{\sum_{k=1}^n (f'_k(t_i) - f'_k(\tau_k^i))^2} \cdot (t_i - t_{i-1}) < \frac{\varepsilon}{2}.
\end{aligned}$$

Junto con (6.5) se tiene que

$$\left| L_P(K) - \int_a^b \varphi(t) dt \right| < \varepsilon,$$

lo que concluye la demostración. ■

Comentamos que si $(f, [a, b])$ y $(g, [c, d])$ son parametrizaciones equivalentes de K (ver Definición 6.2), entonces en virtud de $\varphi'(t) > 0$ se tiene que

$$\begin{aligned}
L(K) &= \int_c^d \sqrt{\sum_{k=1}^n (g'_k(\tau))^2} d\tau = \int_a^b \sqrt{\sum_{k=1}^n (g'_k(\varphi(t)))^2} \varphi'(t) dt = \int_a^b \sqrt{\sum_{k=1}^n (g'_k(\varphi(t))\varphi'(t))^2} dt \\
&= \int_a^b \sqrt{\sum_{k=1}^n (f'_k(t))^2} dt.
\end{aligned}$$

Este resultado ilustra que bajo las presuposiciones del Teorema 6.9 la fórmula para la longitud de arco de K es invariante con respecto a parametrizaciones equivalentes de K .

Para el caso especial de una curva K definida por $x = f_1(t) = t$ e $y = f_2(t) = f(t)$ para una función $f : [a, b] \rightarrow \mathbb{R}$, $f \in C^1[a, b]$ dada, obtenemos

$$(f'_1(t))^2 + (f'_2(t))^2 = 1 + (f'(t))^2$$

y por lo tanto

$$L(K) = \int_a^b \sqrt{1 + (f'(t))^2} dt.$$

Por otro lado, si $K \subset \mathbb{R}^n$ es una curva suave por trozos con la parametrización $(f, [a, b])$, podemos utilizar el vector $\vec{f}'(t) = \{f'_1(t), \dots, f'_n(t)\}$ para escribir su longitud de arco como

$$L(K) = \int_a^b \|\vec{f}'(t)\| dt.$$

Ejemplo 6.7. *Consideremos la espiral*

$$K = \{(x, y, z) \mid x = r \cos t, y = r \sin t, z = ct, t \in [0, 2\pi]\}$$

(ver Figura 6.9). Aquí obtenemos

$$L(K) = \int_0^{2\pi} \sqrt{(x'(t))^2 + (y'(t))^2 + (z'(t))^2} dt = \int_0^{2\pi} \sqrt{r^2 \sin^2 t + r^2 \cos^2 t + c^2} dt$$

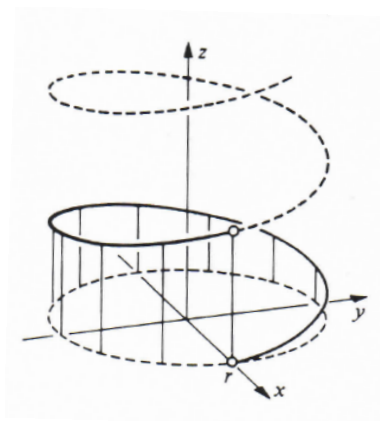


FIGURA 6.9. Ilustración del Ejemplo 6.7

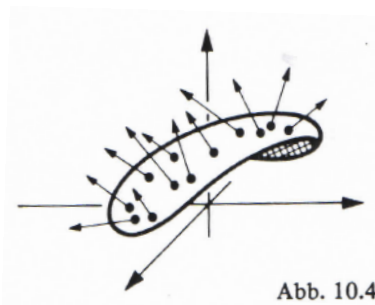


FIGURA 6.10. Un campo vectorial.

$$= 2\pi\sqrt{r^2 + c^2}.$$

6.4. Campos vectoriales; la divergencia y el rotacional

Definición 6.7. Se considera la aplicación $\vec{V} : \mathbb{R}^n \rightarrow \mathbb{V}^n$ con el dominio $D(\vec{V}) = X \subset \mathbb{R}^n$. La totalidad de los vectores $\vec{V}(x) = \{V_1(x), \dots, V_n(x)\}$ se llama campo vectorial sobre X .

Visualizamos campos vectoriales asociando el vector $\vec{V}(x)$ a cada punto $x \in X$ (ver Figura 6.10). En este contexto, nos referimos a una función $f : \mathbb{R}^n \rightarrow \mathbb{R}$ como *función escalar*.

Ejemplo 6.8. El escurrimiento de un fluido por un recipiente genera un campo vectorial si a cada punto del interior del recipiente asociamos el vector de velocidad de una “partícula” del fluido.

Teorema 6.10. Un campo vectorial $\vec{V} : \mathbb{R}^n \rightarrow \mathbb{V}^n$ es continuo sobre $X \subset D(\vec{V})$ si y sólo si cada función de componente V_i ($i = 1, \dots, n$) es continua sobre X .

Demostración. La demostración es análoga al Teorema 1.4. ■

Ejemplo 6.9. Si $K \subset \mathbb{R}^n$ es una curva suave representada por $(f, [a, b])$, entonces la totalidad de sus vectores tangenciales $\vec{T}_f(t)$ con $t \in [a, b]$ forma un campo vectorial sobre K ; a cada punto $x = f(t) = (f_1(t), \dots, f_n(t))$ se asocia el vector

$$\vec{T}_f(t) = \left\{ \frac{f'_1(t)}{\|\vec{f}'(t)\|}, \dots, \frac{f'_n(t)}{\|\vec{f}'(t)\|} \right\}.$$

Ejemplo 6.10. Si la función $f : \mathbb{R}^n \rightarrow \mathbb{R}$ posee todas las derivadas parciales sobre un conjunto abierto X , entonces

$$\text{grad } f(x) = \nabla f(x) = \left\{ \frac{\partial f}{\partial x_1}(x), \dots, \frac{\partial f}{\partial x_n}(x) \right\}$$

define un campo vectorial sobre X .

Teorema 6.11. Sean $f, g : \mathbb{R}^n \rightarrow \mathbb{R}$ y sean $f, g \in C^1(X)$ para un conjunto abierto X . Entonces

1. $\text{grad}(f + g)(x) = \text{grad } f(x) + \text{grad } g(x)$,
2. $\text{grad}(f \cdot g)(x) = f(x) \text{grad } g(x) + g(x) \text{grad } f(x)$.

Demostración. Tarea. ■

Definición 6.8. Sean $\vec{V} = \{V_1, \dots, V_n\} : \mathbb{R}^n \rightarrow \mathbb{V}^n$ un campo vectorial y $X \subset D(\vec{V})$ un conjunto abierto, y sea para cada $\nu = 1, \dots, n$ la función V_ν parcialmente diferenciable con respecto a x_ν sobre X . Entonces la expresión

$$\text{div } \vec{V}(x) = \nabla \cdot \vec{V}(x) = \frac{\partial V_1}{\partial x_1}(x) + \dots + \frac{\partial V_n}{\partial x_n}(x) = \sum_{\nu=1}^n \frac{\partial V_\nu}{\partial x_\nu}(x)$$

se llama la divergencia del campo vectorial $\vec{V}(x)$ sobre X .

La divergencia de un campo vectorial es definida para un número de dimensiones n arbitrario, al contrario del rotacional, que se considera solamente para $n = 3$.

Definición 6.9. Sean $\vec{V} = \{L, M, N\} : \mathbb{R}^3 \rightarrow \mathbb{V}^3$ un campo vectorial y $X \subset D(\vec{V})$ un conjunto abierto, y sean las funciones L , M y N parcialmente diferenciables con respecto a x , y y z sobre X . Entonces la expresión

$$\text{rot } \vec{V}(x, y, z) = \left(\begin{aligned} &\frac{\partial N}{\partial y}(x, y, z) - \frac{\partial M}{\partial z}(x, y, z), \\ &\frac{\partial L}{\partial z}(x, y, z) - \frac{\partial N}{\partial x}(x, y, z), \\ &\frac{\partial M}{\partial x}(x, y, z) - \frac{\partial L}{\partial y}(x, y, z) \end{aligned} \right)$$

se llama el rotacional del campo vectorial $\vec{V}(x)$ sobre X .

Comentamos que si utilizamos

$$\text{grad} = \nabla = \left\{ \frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right\}$$

como vector simbólico, obtenemos

$$\text{rot } \vec{V}(x, y, z) = (\text{grad} \times \vec{V})(x, y, z) = (\nabla \times \vec{V})(x, y, z).$$

Teorema 6.12. Sean $\vec{V}, \vec{W} : \mathbb{R}^3 \rightarrow \mathbb{V}^3$ campos vectoriales, y sobre el conjunto X sean las funciones de componente de $\vec{V}$ y $\vec{W}$ parcialmente diferenciables con respecto a x , y y z . Entonces se tiene que

$$\text{rot}(\vec{V} + \vec{W})(x, y, z) = \text{rot } \vec{V}(x, y, z) + \text{rot } \vec{W}(x, y, z).$$

Demostración. Tarea. ■

Para los campos vectoriales la definición es útil.

Definición 6.10. Se considera el campo vectorial $\vec{V} : \mathbb{R}^n \rightarrow \mathbb{V}^n$, donde $\vec{V} = \{V_1, \dots, V_n\}$. Sea $X \subset \mathbb{R}^n$ un conjunto abierto y $k \in \mathbb{N}_0$. Escribimos $\vec{V} \in C^k(X)$ si $V_\nu \in C^k(X)$ sobre X para todo $\nu = 1, \dots, n$.

Teorema 6.13. Sea $\vec{V} : \mathbb{R}^3 \rightarrow \mathbb{V}$ un campo vectorial, sea $D(\vec{V}) = X$ abierto y $\vec{V} \in C^2(X)$. Entonces se tiene que

$$\text{div}(\text{rot } \vec{V})(x, y, z) = 0 \quad \text{para todo } (x, y, z) \in X.$$

Demostración. Tarea. ■

Teorema 6.14.

1. Sean $\vec{V}, \vec{W} : \mathbb{R}^n \rightarrow \mathbb{V}^n$ campos vectoriales y sea f una función escalar. Sobre un conjunto abierto $X \subset \mathbb{R}^n$ sean $\vec{V}, \vec{W}, f \in C^1(X)$. Entonces se tiene que
 - a) $\text{div}(\vec{V} + \vec{W}) = \text{div } \vec{V} + \text{div } \vec{W}$,
 - b) $\text{div}(f \cdot \vec{V}) = \text{grad } f \cdot \vec{V} + f \cdot \text{div } \vec{V}$.
2. Sean $\vec{V}, \vec{W} : \mathbb{R}^3 \rightarrow \mathbb{V}^3$ campos vectoriales y sea f una función escalar. Sobre un conjunto abierto $X \subset \mathbb{R}^3$ sean $\vec{V}, \vec{W}, f \in C^1(X)$. Entonces se tiene que
 - a) $\text{rot}(f \cdot \vec{V}) = f \cdot \text{rot } \vec{V} - \vec{V} \times \text{grad } f$,
 - b) $\text{div}(\vec{V} \times \vec{W}) = \vec{W} \cdot \text{rot } \vec{V} - \vec{V} \cdot \text{rot } \vec{W}$,
 - c) $\text{rot}(\vec{V} \times \vec{W}) = (\vec{W} \text{ grad})\vec{V} - (\vec{V} \text{ grad})\vec{W} + \vec{V} \text{ div } \vec{W} - \vec{W} \text{ div } \vec{V}$,
 - d) $\text{grad}(\vec{V} \cdot \vec{W}) = (\vec{W} \text{ grad})\vec{V} + (\vec{V} \text{ grad})\vec{W} + \vec{V} \times \text{rot } \vec{W} + \vec{W} \times \text{rot } \vec{V}$.
3. Si además $\vec{V}, f \in C^2(X)$, sabemos que
 - a) $\text{rot}(\text{grad } f) = \vec{0}$,
 - b) $\text{rot}(\text{rot } \vec{V}) = \text{grad}(\text{div } \vec{V}) - \Delta \vec{V}$, donde para $\vec{V} = \{V_1, V_2, V_3\}$ se define

$$\Delta \vec{V} := \{\Delta V_1, \Delta V_2, \Delta V_3\}.$$

Demostración. Tarea. ■

6.5. Integrales de línea

Para motivar el concepto de la integral de línea, consideremos un problema de física. Supongamos que un cuerpo se traslada a lo largo de una curva K bajo la influencia de un campo de fuerza $\vec{V}$. Describimos la curva K en dependencia del tiempo t . Sea $(f, [a, b])$ una parametrización de K . Para calcular el trabajo definimos una partición $P = \{t_0, \dots, t_m\}$ del intervalo $[a, b]$. El trabajo que corresponde al periodo $[t_{k-1}, t_k]$ es aproximadamente

$$[\vec{V}(f(t_{k-1})) \cdot \vec{f}'(t_{k-1})](t_k - t_{k-1}),$$

puesto que aquí importa solamente la componente de la fuerza en la dirección del movimiento. Entonces, la fuerza total es aproximadamente igual a la suma de Riemann

$$\sum_{k=1}^m [\vec{V}(f(t_{k-1})) \cdot \vec{f}'(t_{k-1})](t_k - t_{k-1}),$$

la cual, bajo hipótesis apropiadas, convergerá a la integral

$$\int_a^b \vec{V}(f(t)) \cdot \vec{f}'(t) dt.$$

Esta consideración motiva la siguiente definición.

Definición 6.11. Sea K una curva suave por trozos en $\mathbb{R}^n$ con la parametrización $(f, [a, b])$, además sea $\vec{V} = \{V_1, \dots, V_n\}$ un campo vectorial continuo sobre K . Entonces la expresión

$$\int_a^b \vec{V}(f(t)) \cdot \vec{f}'(t) dt = \int_K (V_1 dx_1 + \dots + V_n dx_n) \quad (6.6)$$

se llama integral de línea de $\vec{V}$ a lo largo de K con respecto a la parametrización $(f, [a, b])$.

Comentamos que la integral

$$\int_a^b \vec{V}(f(t)) \cdot \vec{f}'(t) dt$$

existe puesto que el integrando es continuo con la excepción de a lo más un número finito de puntos.

Por otro lado, sea $(g, [c, d])$ una parametrización equivalente de K . Según la Definición 6.2, se tiene lo siguiente con $\tau = \varphi(t)$:

$$\int_c^d \vec{V}(g(\tau)) \cdot \vec{g}'(\tau) d\tau = \int_a^b \vec{V}(g(\varphi(t))) \cdot \vec{g}'(\varphi(t)) \cdot \varphi'(t) dt = \int_a^b \vec{V}(f(t)) \cdot \vec{f}'(t) dt,$$

es decir, el valor de la integral no cambia si pasamos a parametrizaciones equivalentes de la curva K .

Finalmente, justificamos la notación informal, pero muy útil, en (6.6). Para tal efecto, considerando

$$\begin{aligned} \vec{V} &= \vec{V}(x_1, \dots, x_n) = \{V_1(x_1, \dots, x_n), \dots, V_n(x_1, \dots, x_n)\}, \\ f(t) &= (f_1(t), f_2(t), \dots, f_n(t)), \\ \vec{f}'(t) &= \{f'_1(t), f'_2(t), \dots, f'_n(t)\} \end{aligned}$$

obtenemos

$$\vec{V}(f(t)) = \{V_1(f_1(t), \dots, f_n(t)), \dots, V_n(f_1(t), \dots, f_n(t))\},$$

y $\vec{V}(f(t)) \cdot \vec{f}'(t)$ es el producto escalar

$$\vec{V}(f(t)) \cdot \vec{f}'(t) = \sum_{i=1}^n V_i(f_1(t), \dots, f_n(t)) \cdot f'_i(t),$$

por lo tanto

$$\begin{aligned} \int_a^b \vec{V}(f(t)) \cdot \vec{f}'(t) dt &= \int_a^b \left(\sum_{i=1}^n V_i(f_1(t), \dots, f_n(t)) \cdot f'_i(t) \right) dt \\ &= \sum_{i=1}^n \left(\int_a^b V_i(f_1(t), \dots, f_n(t)) \cdot f'_i(t) dt \right). \end{aligned}$$

si ponemos $x_i = f_i(t)$, entonces

$$\frac{dx_i}{dt} = f'_i(t)$$

o simbólicamente $f'_i(t) dt = dx_i$, lo que justifica la notación (6.6).

Ejemplo 6.11. *Se considera el campo vectorial*

$$\vec{V}(x, y, z) = \{yz, xz, xy\},$$

y queremos calcular la integral de línea

$$\int_K (yz dx + xz dy + xy dz)$$

para dos curvas diferentes.

1. Sea K_1 la curva con la parametrización

$$x = t, \quad y = t^2, \quad z = 2, \quad t \in [0, 1].$$

La curva conecta en $\mathbb{R}^3$ los puntos $(0, 0, 2)$ y $(1, 1, 2)$ a lo largo de una parábola. Aquí tenemos que

$$\vec{f}'(t) = \{1, 2t, 0\}$$

y por lo tanto

$$\begin{aligned} \int_{K_1} (yz dx + xz dy + xy dz) &= \int_0^1 \{2t^2, 2t, t^3\} \cdot \{1, 2t, 0\} dt \\ &= \int_0^1 6t^2 dt = [2t^3]_{t=0}^{t=1} = 2. \end{aligned}$$

2. Sea ahora K_2 la curva con la parametrización

$$x = t, \quad y = t, \quad z = 2, \quad t \in [0, 1].$$

La curva conecta en $\mathbb{R}^3$ los mismos puntos $(0, 0, 2)$ y $(1, 1, 2)$ a lo largo de una recta. Aquí obtenemos

$$\vec{f}'(t) = \{1, 1, 0\}$$

y por lo tanto

$$\begin{aligned} \int_{K_2} (yz \, dx + xz \, dy + xy \, dz) &= \int_0^1 \{2t, 2t, t^2\} \cdot \{1, 1, 0\} \, dt \\ &= \int_0^1 4t \, dt = [2t^2]_{t=0}^{t=1} = 2, \end{aligned}$$

es decir aquí ambas integrales de línea tienen el mismo valor.

Ejemplo 6.12. Consideramos el campo vectorial

$$\vec{V}(x, y, z) = \{x^2y, x - z, xyz\},$$

y queremos calcular las integrales de línea a lo largo de las curvas K_1 y K_2 del Ejemplo 6.11.

1. Para la integración a lo largo de K_1 obtenemos

$$\begin{aligned} \int_{K_1} (x^2y \, dx + (x - z) \, dy + xyz \, dz) &= \int_0^1 \{t^4, t - 2, 2t^3\} \cdot \{1, 2t, 0\} \, dt \\ &= \int_0^1 (t^4 + 2t^2 - 4t) \, dt = -\frac{17}{15}. \end{aligned}$$

2. Para la integración a lo largo de K_2 obtenemos

$$\begin{aligned} \int_{K_2} (x^2y \, dx + (x - z) \, dy + xyz \, dz) &= \int_0^1 \{t^3, t - 2, 2t^2\} \cdot \{1, 1, 0\} \, dt \\ &= \int_0^1 (t^3 + t - 2) \, dt = -\frac{5}{4}. \end{aligned}$$

Este ejemplo demuestra que en general la integral de línea es “dependiente del camino”, es decir que si dos curvas K_1 y K_2 conectan los mismos puntos, las integrales de línea a lo largo de ambas curvas (con el mismo campo vectorial) pueden tener valores diferentes.

6.6. Potenciales e independencia del camino de integrales de línea

En virtud de los resultados de la sección anterior parece deseable caracterizar aquellos campos vectoriales que entregan integrales de línea independientes del camino. Resultará que de gran importancia son aquellos campos vectoriales que pueden ser representados como gradiente de una función escalar.

Definición 6.12. Sea $\vec{V} : \mathbb{R}^n \rightarrow \mathbb{V}^n$ un campo vectorial, y sea $D(\vec{V}) = X$ abierto. Si existe una función escalar $\varphi \in C^1(X)$ tal que

$$\vec{V}(x) = \text{grad } \varphi(x), \quad x \in X,$$

entonces φ se llama potencial de $\vec{V}$. El campo vectorial $\vec{V}$ se llama campo vectorial conservativo.

Ejemplo 6.13. El campo vectorial $\vec{V}(x, y, z) = \{yz, xz, zy\}$ es conservativo porque para $\varphi(x, y, z) = xyz$ se tiene que

$$\text{grad } \varphi(x, y, z) = \{yz, xz, xy\} = \vec{V}(x, y, z).$$

Ejemplo 6.14. El campo vectorial $\vec{V}(x, y, z) = \{x^2y, x - z, xyz\}$ no es conservativo. Si existiera sobre un conjunto abierto $X \subset \mathbb{R}^3$ con una función $\varphi \in C^1(X)$ tal que

$$\text{grad } \varphi(x, y, z) = \vec{V}(x, y, z),$$

entonces, debido a

$$\varphi_x(x, y, z) = x^2y, \quad \varphi_y(x, y, z) = x - z, \quad \varphi_z(x, y, z) = xyz$$

se debería tener que $\varphi \in C^2(X)$ y por lo tanto $\varphi_{xy} = \varphi_{yx}$, lo que evidentemente no es válido.

Ejemplo 6.15. Para $x > 0$ e $y > 0$ sea un campo vectorial definido por

$$\vec{V}(x, y) = \left\{ -\frac{y}{x^2 + y^2}, \frac{x}{x^2 + y^2} \right\}.$$

Observamos que la función $\varphi(x, y) = \arctan(y/x)$ satisface

$$\varphi_x(x, y) = -\frac{y}{x^2 + y^2}, \quad \varphi_y(x, y) = \frac{x}{x^2 + y^2},$$

es decir, $\vec{V}(x, y) = \text{grad } \varphi(x, y)$, por lo tanto $\vec{V}$ es un campo vectorial conservativo y φ es un potencial de $\vec{V}$.

Los campos vectoriales que poseen un potencial tienen una propiedad muy interesante con respecto a integrales de línea.

Definición 6.13. Sea $\vec{V} : \mathbb{R}^n \rightarrow \mathbb{V}^n$ un campo vectorial, sea $D(\vec{V}) = X$ abierto y sea $\vec{V}$ continuo sobre X . Sean P_1 y P_2 puntos arbitrarios de X . Si la integral de línea

$$\int_{K(P_1, P_2)} \{V_1 dx_1 + \cdots + V_n dx_n\}$$

posee el mismo valor para todas las curvas suaves por trozos $K(P_1, P_2) \subset X$ que conecten P_1 con P_2 , entonces la integral de línea se llama independiente del camino con respecto al campo vectorial $\vec{V}$.

Ahora podemos formular el siguiente teorema.

Teorema 6.15. Sea $\vec{V} = (V_1, \dots, V_n) : \mathbb{R}^n \rightarrow \mathbb{V}^n$ un campo vectorial, sea $X = D(\vec{V})$ un conjunto abierto y convexo, y sea $\vec{V}$ continuo sobre X . Entonces los siguientes enunciados son equivalentes:

1. *El campo vectorial $\vec{V}$ posee un potencial.*
2. *La integral de línea con respecto a $\vec{V}$ es independiente del camino.*

Demostración.

1. Supongamos que el campo vectorial $\vec{V}$ posee un potencial denotado φ , es decir se tiene que $\vec{V} = \text{grad } \varphi$. Sean P_1 y P_2 puntos arbitrarios de X , y sea K una curva suave por trozos que conecta P_1 con P_2 . Sea $(f, [a, b])$ una parametrización de K . Entonces

$$\begin{aligned} \int_K (V_1 dx_1 + \cdots + V_n dx_n) &= \int_a^b \vec{V}(f(t)) \cdot \vec{f}'(t) dt \\ &= \int_a^b \sum_{i=1}^n \left(\frac{\partial \varphi}{\partial x_i}(f(t)) \frac{df_i}{dt}(t) \right) dt \\ &= \varphi(f(b)) - \varphi(f(a)) = \varphi(P_2) - \varphi(P_1), \end{aligned}$$

por lo tanto la integral de línea es independiente del camino.

2. Supongamos ahora que la integral de línea es independiente del camino. Sea $P_0 \in X$ un punto fijo. Definimos sobre X la función $\varphi : \mathbb{R}^n \rightarrow \mathbb{R}$ dada por

$$\varphi(x) := \int_{P_0}^x \vec{V}(f(t)) \cdot \vec{f}'(t) dt,$$

donde elegimos alguna curva suave que conecta P_0 con x . Entonces, para cada $h > 0$ suficientemente pequeño, $x = (x_1, \dots, x_n)$ y cada $i = 1, \dots, n$ sabemos que

$$\begin{aligned} D_i(h) &:= \frac{1}{h} (\varphi(x_1, \dots, x_{i-1}, x_i + h, x_{i+1}, \dots, x_n) \\ &\quad - \varphi(x_1, \dots, x_{i-1}, x_i, x_{i+1}, \dots, x_n)) \\ &= \frac{1}{h} \int_0^h V_i(x_1, \dots, x_{i-1}, x_i + t, x_{i+1}, \dots, x_n) dt, \end{aligned}$$

donde la curva de integración es el segmento lineal entre $(x_1, \dots, x_i + h, \dots, x_n)$ y $(x_1, \dots, x_{i-1}, x_i, x_{i+1}, \dots, x_n)$ que es completamente contenido en X debido a la convexidad de X . Según el Teorema del Valor Intermedio del Cálculo Integral existe un ξ entre 0 y h tal que

$$D_i(h) = \frac{1}{h} \cdot V_i(x_1, \dots, x_{i-1}, x_i + \xi, x_{i+1}, \dots, x_n) \cdot h,$$

luego

$$\begin{aligned} \frac{\partial \varphi}{\partial x_i}(x) &= \lim_{h \rightarrow 0} D_i(h) = \lim_{h \rightarrow 0} V_i(x_1, \dots, x_{i-1}, x_i + \xi, x_{i+1}, \dots, x_n) \\ &= V_i(x_1, \dots, x_{i-1}, x_i, x_{i+1}, \dots, x_n) = V_i(x), \end{aligned}$$

por lo tanto φ es un potencial de $\vec{V}$. ■

Comentamos que este teorema entrega las siguientes informaciones importantes. Si para un campo vectorial continuo $\vec{V}$ se conoce un potencial φ , podemos calcular las integrales de línea de manera muy fácil, puesto que si P_1 y P_2 son puntos en X , se tiene para cada curva suave por trozos $K(P_1, P_2)$ que conecta P_1 con P_2 que

$$\int_{K(P_1, P_2)} (V_1 dx_1 + \cdots + V_n dx_n) = \varphi(P_2) - \varphi(P_1).$$

Por otro lado, si $\vec{V}$ es un campo vectorial para el cual todas las integrales de línea

$$\int_{K(P_1, P_2)} (V_1 dx_1 + \cdots + V_n dx_n), \quad P_1, P_2 \in X$$

son iguales, un potencial de $\vec{V}$ puede ser determinado por

$$\varphi(x) = \int_{P_0}^x \vec{V}(f(t)) \cdot \vec{f}'(t) dt,$$

donde se integra sobre alguna curva suave que conecta el punto fijo P_0 con el punto variable x .

Revisando la demostración del Teorema 6.15 vemos que la restricción a dominios X convexos es muy restrictiva, y que efectivamente mediante una ligera modificación de la demostración podemos probar este teorema queda válido para dominios X más generales, por ejemplos para dominios que *tienen forma de estrella*.

Definición 6.14. Sea $X \subset \mathbb{R}^n$ un conjunto abierto. Se dice que X tiene forma de estrella con respecto al punto $x_0 \in X$ si para cada punto $x \in X$ el segmento lineal $\overline{x_0, x}$ pertenece enteramente a X .

Evidentemente, todo conjunto X convexo tiene forma de estrella.

El siguiente teorema formula un criterio mediante el cual podemos verificar si un campo vectorial $\vec{V}(x)$ posee un potencial.

Teorema 6.16. Supongamos que el conjunto abierto $X \subset \mathbb{R}^n$ tiene forma de estrella, y que el campo vectorial $\vec{V} = (V_1, \dots, V_n) : X \rightarrow \mathbb{R}^n$ satisface la siguiente condición de integrabilidad:

$$\frac{\partial V_i}{\partial x_j} = \frac{\partial V_j}{\partial x_i} \quad \text{en } X \text{ para } i, j = 1, \dots, n. \quad (6.7)$$

En este caso el campo vectorial $\vec{V}$ posee un potencial φ .

Demostración. Aplicando una translación si necesario podemos lograr que $x_0 = 0$; en lo siguiente supondremos esto y definiremos $\varphi(x)$ como integral de línea a lo largo del segmento lineal $\phi(t) = tx$, $0 \leq t \leq 1$, de 0 a x para $x \in X$:

$$\varphi(x) := \int_0^x (V_1 dy_1 + \cdots + V_n dy_n) = \int_0^1 \vec{V}(tx) \cdot x dt. \quad (6.8)$$

Ahora tomamos en cuenta que

$$\vec{V}(tx) \cdot x = V_1(tx)x_1 + \cdots + V_n(tx)x_n,$$

es decir

$$\frac{\partial}{\partial x_1}(\vec{V}(tx) \cdot x) = V_1(tx) + \sum_{i=1}^n x_i t \frac{\partial V_i}{\partial x_1}(tx) = V_1(tx) + (\text{grad } V_1(tx)) \cdot (tx); \quad (6.9)$$

la última identidad es válida debido a que en virtud de (6.7), $\partial V_i / \partial x_1 = \partial V_1 / \partial x_i$. Por otro lado,

$$\frac{d}{dt}(tV_1(tx)) = V_1(tx) + t \sum_{i=1}^n x_i \frac{\partial V_1}{\partial x_i}(tx). \quad (6.10)$$

Notamos que las expresiones en los lados derechos de (6.9) y (6.10) son idénticas. Ahora calculamos que

$$\frac{\partial}{\partial x_1} \varphi(x) = \int_0^1 \frac{\partial}{\partial x_1}(\vec{V}(tx) \cdot x) dt = \int_0^1 \frac{d}{dt}(tV_1(tx)) dt = [tV_1(tx)]_{t=0}^{t=1} = V_1(x).$$

De la misma forma evaluamos las demás derivadas parciales, y terminamos con la fórmula $\text{grad } \varphi = \vec{V}$ sobre X . ■

Ejemplo 6.16. Queremos calcular la integral de línea

$$I = \int_K (2xy dx + (x^2 + y^2) dy),$$

donde K es la parte de la elipse $9x^2 + 4y^2 = 36$ localizada en el cuadrante $x \geq 0, y \geq 0$. Esta curva conecta el punto $(2, 0)$ con el punto $(0, 3)$. En este caso,

$$V_1(x, y) = 2xy, \quad V_2(x, y) = x^2 + y^2.$$

Aquí calculamos

$$\frac{\partial V_1}{\partial y} = 2x, \quad \frac{\partial V_2}{\partial x} = 2x,$$

dado que ambas expresiones coinciden en todas partes, en particular sobre un dominio en forma de estrella con respecto a $x_0 = (0, 0)$, el campo vectorial $\vec{V}(x, y) = \{V_1(x, y), V_2(x, y)\}$ es conservativo y podemos calcular su potencial φ mediante la fórmula (6.8), es decir

$$\begin{aligned} \varphi(x, y) &= \int_0^1 \vec{V}(tx, ty) \cdot (x, y) dt = \int_0^1 (2t^2 xy \cdot x + t^2(x^2 + y^2) \cdot y) dt \\ &= (3x^2 y + y^3) \int_0^1 t^2 dt = (3x^2 y + y^3) \cdot \left[\frac{1}{3} t^3 \right]_{t=0}^{t=1} = x^2 y + \frac{1}{3} y^3, \end{aligned}$$

es decir

$$\vec{V}(x, y) = \text{grad } \varphi(x, y) = \text{grad} \left(x^2 y + \frac{1}{3} y^3 \right)$$

y por lo tanto

$$I = \varphi(0, 3) - \varphi(2, 0) = 9 - 0 = 9.$$

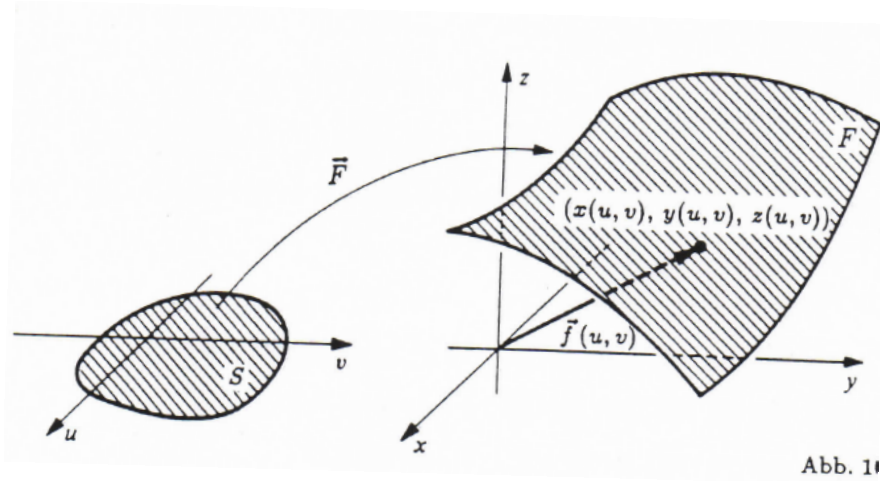


FIGURA 6.11. Ilustración de la Definición 6.15.

Ejemplo 6.17. Queremos calcular la integral de línea

$$I = \int_K (yz \, dx + xz \, dy + xy \, dz),$$

donde K es alguna curva suave por trozos que conecta los puntos $(0, 0, 2)$ y $(1, 1, 2)$ (ver Ejemplo 6.11). Para verificar que esta integral efectivamente es independiente del camino, verificamos primeramente si el campo vectorial $\vec{V}(x, y, z) = (V_1, V_2, V_3)(x, y, z)$ con $V_1(x, y, z) = yz$, $V_2(x, y, z) = xz$, $V_3(x, y, z) = xy$ es conservativo. Para tal efecto verifiquemos la satisfacción de las condiciones de integrabilidad

$$\frac{\partial V_1}{\partial y} = \frac{\partial V_2}{\partial x}, \quad \frac{\partial V_1}{\partial z} = \frac{\partial V_3}{\partial x}, \quad \frac{\partial V_2}{\partial z} = \frac{\partial V_3}{\partial y};$$

efectivamente, aquí las tres igualdades están satisfechas, por lo tanto $\vec{V}$ posee un potencial, el cual puede ser determinado de manera similar al ejemplo anterior:

$$\begin{aligned} \varphi(x, y, z) &= \int_0^1 \vec{V}(tx, ty, tz) \cdot (x, y, z) \, dt = \int_0^1 (t^2 yz \cdot x + t^2 xz \cdot y + t^2 xy \cdot z) \, dt \\ &= 3xyz \int_0^1 t^2 \, dt = 3xyz \cdot \left[\frac{1}{3} t^3 \right]_{t=0}^{t=1} = xyz \end{aligned}$$

es decir

$$\vec{V}(x, y, z) = \text{grad } \varphi(x, y, z) = \text{grad } xyz$$

y por lo tanto

$$I = \varphi(1, 1, 2) - \varphi(0, 0, 2) = 2.$$

6.7. Superficies en $\mathbb{R}^3$

Definición 6.15. Sea $f = (x(u, v), y(u, v), z(u, v)) : \mathbb{R}^2 \rightarrow \mathbb{R}^3$, sea $D(f) = S$ abierto, sea $f \in C^1(S)$, y sea el rango de la matriz

$$\frac{\partial(x, y, z)}{\partial(u, v)} = \begin{bmatrix} \frac{\partial x}{\partial u} & \frac{\partial y}{\partial u} & \frac{\partial z}{\partial u} \\ \frac{\partial x}{\partial v} & \frac{\partial y}{\partial v} & \frac{\partial z}{\partial v} \end{bmatrix}$$

igual 2. Entonces el conjunto

$$F = \{(x, y, z) \mid x = x(u, v), y = y(u, v), z = z(u, v), (u, v) \in S\}$$

se llama superficie en $\mathbb{R}^3$, y (f, S) es una parametrización de F , ver Figura 6.11.

En muchas situaciones también representamos una superficie S vectorialmente, es decir representamos F por

$$\vec{f}(u, v) = \{x(u, v), y(u, v), z(u, v)\}, \quad (u, v) \in S.$$

Ejemplo 6.18. Sea F dada por

$$x(u, v) = u, \quad y(u, v) = v, \quad z(u, v) = 0, \quad (u, v) \in S = \mathbb{R}^2.$$

Entonces sobre S se tiene que

$$\frac{\partial(x, y, z)}{\partial(u, v)} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}.$$

El rango de esta matriz es 2, por lo tanto F representa una superficie en $\mathbb{R}^3$; se trata del plano (x, y) .

Ejemplo 6.19. Sea F dada por

$$\begin{aligned} x(u, v) &= u \cos v, & y(u, v) &= u \sin v, & z(u, v) &= 0, \\ S &= \{(u, v) \mid 0 < u < \infty, 0 < v < 2\pi\}. \end{aligned}$$

Entonces sobre S se tiene que

$$\frac{\partial(x, y, z)}{\partial(u, v)} = \begin{bmatrix} \cos v & \sin v & 0 \\ -u \sin v & u \cos v & 0 \end{bmatrix}.$$

El rango de esta matriz es 2, por lo tanto F representa una superficie en $\mathbb{R}^3$.

Ejemplo 6.20. Sea F dada por

$$x(u, v) = u, \quad y(u, v) = v, \quad z(u, v) = \sqrt{R^2 - u^2 - v^2}, \quad S = \{(u, v) \mid u^2 + v^2 < R^2\}.$$

Entonces sobre S se tiene que

$$\frac{\partial(x, y, z)}{\partial(u, v)} = \begin{bmatrix} 1 & 0 & -\frac{u}{\sqrt{R^2 - u^2 - v^2}} \\ 0 & 1 & -\frac{v}{\sqrt{R^2 - u^2 - v^2}} \end{bmatrix}.$$

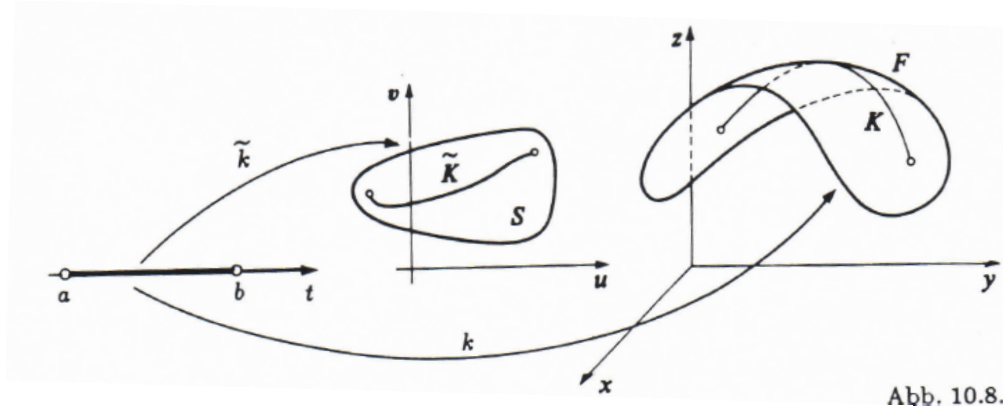


FIGURA 6.12. Ilustración del Teorema 6.17.

El rango de esta matriz es 2, por lo tanto F representa una superficie en $\mathbb{R}^3$; efectivamente, se trata del hemisferio superior de una esfera con el centro $(0,0,0)$ y el radio R .

6.8. Curvas en superficies, planos tangenciales y vectores normales

Teorema 6.17. Se considera una superficie F en $\mathbb{R}^3$ con la parametrización (f, S) ; además sea $\tilde{K}$ una curva suave en $\mathbb{R}^2$ que pertenece enteramente a S . Sea una parametrización de $\tilde{K}$ dada por $(\tilde{k}, [a, b])$ con $\tilde{k}(t) = (\tilde{u}(t), \tilde{v}(t))$. Entonces a través de $k(t) = f(\tilde{k}(t))$, $t \in [a, b]$, está definida una curva suave K en $\mathbb{R}^3$ que enteramente pertenece a la superficie F , ver Figura 6.12.

Demostración. Es evidente que K es una curva que completamente pertenece a F . Solamente queda para demostrar que K es una curva suave. Para tal efecto calculamos para un valor del parámetro $t_0 \in [a, b]$ el vector tangencial en el punto $k(t_0)$ de la curva. Utilizando la regla de la cadena y definiendo $u_0 := \tilde{u}(t_0)$, $v_0 := \tilde{v}(t_0)$ obtenemos

$$\vec{k}'(t_0) = \left\{ \begin{aligned} &\frac{\partial x}{\partial u}(u_0, v_0) \cdot \tilde{u}'(t_0) + \frac{\partial x}{\partial v}(u_0, v_0) \cdot \tilde{v}'(t_0), \\ &\frac{\partial y}{\partial u}(u_0, v_0) \cdot \tilde{u}'(t_0) + \frac{\partial y}{\partial v}(u_0, v_0) \cdot \tilde{v}'(t_0), \\ &\frac{\partial z}{\partial u}(u_0, v_0) \cdot \tilde{u}'(t_0) + \frac{\partial z}{\partial v}(u_0, v_0) \cdot \tilde{v}'(t_0) \end{aligned} \right\}.$$

Definiendo

$$\vec{f}_u := \left\{ \frac{\partial x}{\partial u}, \frac{\partial y}{\partial u}, \frac{\partial z}{\partial u} \right\}, \quad \vec{f}_v := \left\{ \frac{\partial x}{\partial v}, \frac{\partial y}{\partial v}, \frac{\partial z}{\partial v} \right\},$$

obtenemos

$$\vec{k}'(t_0) = \vec{f}_u(u_0, v_0) \cdot \tilde{u}'(t_0) + \vec{f}_v(u_0, v_0) \cdot \tilde{v}'(t_0).$$

Puesto que

$$\vec{k}'(t_0) = \{\tilde{u}'(t_0), \tilde{v}'(t_0)\} \neq 0,$$

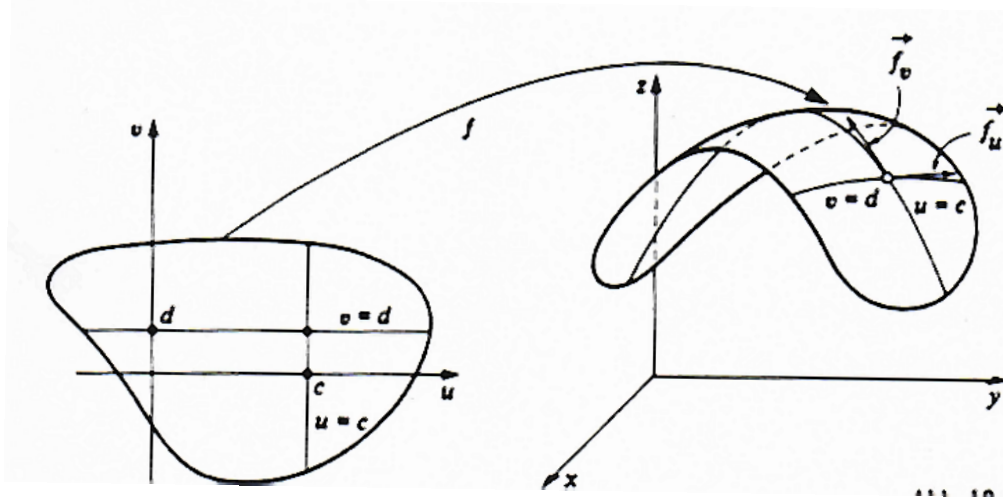


FIGURA 6.13. Los vectores $\vec{f}_u$ y $\vec{f}_v$ como vectores tangenciales de las líneas de parámetros.

la independencia lineal de los vectores $\vec{f}_u(u_0, v_0)$ y $\vec{f}_v(u_0, v_0)$ implica que $\vec{k}'(t_0) \neq \vec{0}$. ■

Comentamos que los vectores linealmente independientes que aparecen en esta demostración,

$$\begin{aligned}\vec{f}_u(u_0, v_0) &= \left\{ \frac{\partial x}{\partial u}(u_0, v_0), \frac{\partial y}{\partial u}(u_0, v_0), \frac{\partial z}{\partial u}(u_0, v_0) \right\}, \\ \vec{f}_v(u_0, v_0) &= \left\{ \frac{\partial x}{\partial v}(u_0, v_0), \frac{\partial y}{\partial v}(u_0, v_0), \frac{\partial z}{\partial v}(u_0, v_0) \right\},\end{aligned}$$

también jugaran un rol importante más adelante. Por mientras, notamos que cada vector tangencial de una curva suave por el punto de la superficie $f(u_0, v_0)$ puede ser representado como una combinación lineal de estos dos vectores.

A su vez, los vectores $\vec{f}_u(u_0, v_0)$ y $\vec{f}_v(u_0, v_0)$ son vectores tangenciales de curvas sobre esta superficie. Considerando las llamadas *líneas de parámetros* $\tilde{k}(t) = (t, d)$ y $\tilde{k}(t) = (c, t)$, donde c y d son constantes, obtenemos los vectores tangenciales asociados $\tilde{k}'(t_0) = \vec{f}_u(u_0, v_0) \cdot 1$ y $\tilde{k}'(t_0) = \vec{f}_v(u_0, v_0) \cdot 1$, respectivamente, ver Figura 6.13.

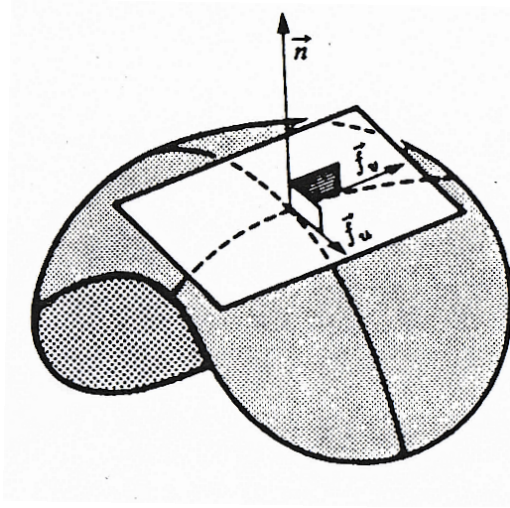
Definición 6.16. Sea F una superficie en $\mathbb{R}^3$ con la parametrización (f, S) , y sea $f(u_0, v_0)$ un punto de F . El plano abierto por los vectores $\vec{f}_u(u_0, v_0)$ y $\vec{f}_v(u_0, v_0)$ y que contiene el punto $f(u_0, v_0)$ se llama plano tangencial de F en el punto $f(u_0, v_0)$.

El vector normal del plano tangencial en un punto es de gran importancia.

Definición 6.17. Sea F una superficie en $\mathbb{R}^3$ con la parametrización (f, S) , y sea $f(u, v)$ un punto de F . Entonces el vector

$$\vec{n}(u, v) = \frac{\vec{f}_u(u, v) \times \vec{f}_v(u, v)}{\|\vec{f}_u(u, v) \times \vec{f}_v(u, v)\|} \quad (6.11)$$

se llama vector normal de F en el punto de superficie $f(u, v)$.

FIGURA 6.14. El vector normal $\vec{n}(u, v)$ dado por (6.11).

Evidentemente, el vector normal $\vec{n}(u, v)$ es un vector normalizado (de la longitud 1) que es ortogonal tanto a $\vec{f}_u(u, v)$ como a $\vec{f}_v(u, v)$. Puesto que cada vector tangencial de una curva suave sobre F por el punto de superficie $f(u, v)$ puede ser representado por una combinación lineal de estos dos vectores, se tiene que $\vec{n}(u, v)$ es ortogonal a cada curva suave que pasa por este punto. En particular, $\vec{n}(u, v)$ es ortogonal al plano tangencial en el punto $f(u, v)$, ver Figura 6.14.

Queremos calcular las componentes del vector $\vec{n}(u, v)$, partiendo de la matriz funcional

$$\frac{\partial(x, y, z)}{\partial(u, v)} = \begin{bmatrix} \frac{\partial x}{\partial u} & \frac{\partial y}{\partial u} & \frac{\partial z}{\partial u} \\ \frac{\partial x}{\partial v} & \frac{\partial y}{\partial v} & \frac{\partial z}{\partial v} \end{bmatrix}. \quad (6.12)$$

Utilizaremos la notación

$$D_x := \det \frac{\partial(y, z)}{\partial(u, v)} = \begin{vmatrix} \frac{\partial y}{\partial u} & \frac{\partial z}{\partial u} \\ \frac{\partial y}{\partial v} & \frac{\partial z}{\partial v} \end{vmatrix}, \quad D_z := \det \frac{\partial(x, y)}{\partial(u, v)} = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial y}{\partial u} \\ \frac{\partial x}{\partial v} & \frac{\partial y}{\partial v} \end{vmatrix},$$

$$D_y := -\det \frac{\partial(x, z)}{\partial(u, v)} = -\begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial z}{\partial u} \\ \frac{\partial x}{\partial v} & \frac{\partial z}{\partial v} \end{vmatrix}, \quad D := \sqrt{D_x^2 + D_y^2 + D_z^2}.$$

Según la Definición 6.15, la matriz (6.12) posee el rango 2, por lo tanto no todos los determinantes D_x , D_y y D_z pueden tener el valor cero, y se debe tener que $D(u, v) > 0$ para todo

$(u, v) \in S$. Según la definición del producto vectorial obtenemos ahora

$$\begin{aligned}\vec{n}(u, v) &= \frac{\vec{f}_u(u, v) \times \vec{f}_v(u, v)}{\|\vec{f}_u(u, v) \times \vec{f}_v(u, v)\|} = \frac{\{D_x(u, v), D_y(u, v), D_z(u, v)\}}{D(u, v)} \\ &= \{\cos \alpha(u, v), \cos \beta(u, v), \cos \gamma(u, v)\},\end{aligned}$$

donde definimos

$$\cos \alpha(u, v) = \frac{D_x(u, v)}{D(u, v)}, \quad \cos \beta(u, v) = \frac{D_y(u, v)}{D(u, v)}, \quad \cos \gamma(u, v) = \frac{D_z(u, v)}{D(u, v)}.$$

En el caso particular $x = u$, $y = v$, es decir, en el caso de una superficie F dada en la forma explícita $z = z(x, y)$ para $(x, y) \in S$ obtenemos

$$D_x = \begin{vmatrix} 0 & \frac{\partial z}{\partial x} \\ 1 & \frac{\partial z}{\partial y} \end{vmatrix} = -\frac{\partial z}{\partial x}, \quad D_y = -\begin{vmatrix} 1 & \frac{\partial z}{\partial x} \\ 0 & \frac{\partial z}{\partial y} \end{vmatrix} = -\frac{\partial z}{\partial y}, \quad D_z = \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = 1$$

y por lo tanto el vector normal

$$\vec{n}(x, y) = \frac{\left\{ -\frac{\partial z}{\partial x}(x, y), -\frac{\partial z}{\partial y}(x, y), 1 \right\}}{\sqrt{\left(\frac{\partial z}{\partial x}(x, y) \right)^2 + \left(\frac{\partial z}{\partial y}(x, y) \right)^2 + 1}}.$$

Puesto que

$$\vec{n}(x, y) = \frac{\vec{f}_x(x, y) \times \vec{f}_y(x, y)}{\|\vec{f}_x(x, y) \times \vec{f}_y(x, y)\|},$$

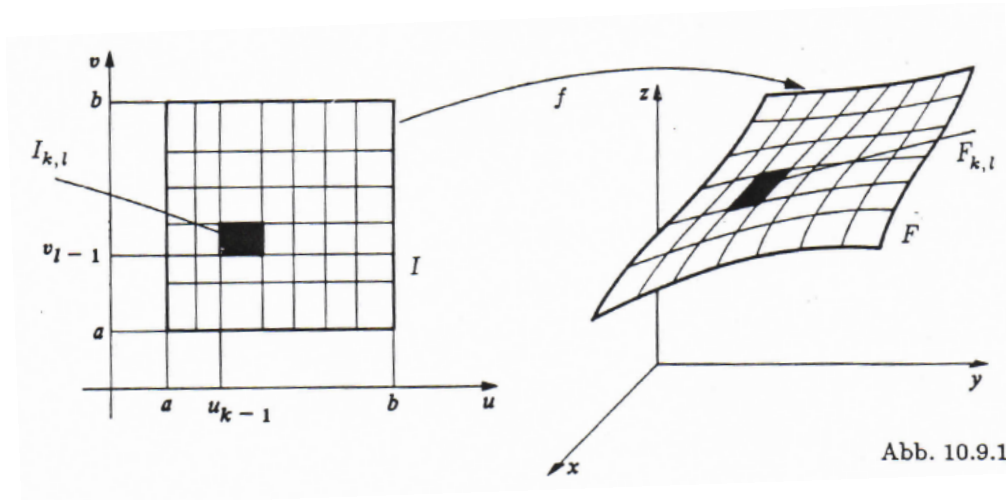
el vector normal siempre apunta hacia arriba; además sabemos que

$$\cos \gamma(x, y) = \frac{1}{\sqrt{\left(\frac{\partial z}{\partial x}(x, y) \right)^2 + \left(\frac{\partial z}{\partial y}(x, y) \right)^2 + 1}}.$$

6.9. El contenido de superficies e integrales de superficie

Es obvio que por lo menos algunas superficies poseen un “contenido”, pero según nuestras experiencias previas, este “contenido” no existe a priori, sino que debe ser definido. Evidentemente, tal definición debe ser más complicada que el contenido n -dimensional de conjuntos Riemann-medibles en $\mathbb{R}^n$, puesto que ahora queremos definir el contenido de un conjunto de puntos de una dimensión inferior el cual está inmerso en un espacio de dimensión mayor.

Al parecer existe una cierta analogía con la situación de la definición de la longitud de arco de una curva, donde definimos la longitud de una curva como el supremo de las longitudes de trazados poligonales, y así reducimos la longitud de una curva a las longitudes de curvas elementales (segmentos lineales). Entonces, parece lógico tratar de aproximar el “contenido superficial” por superficies elementales (por ejemplo, triángulos o rectángulos),

FIGURA 6.15. Partición de I y descomposición de F en trozos de superficie.

y definir el contenido de una superficie como el supremo de las superficies de todas las superficies poliédricas aproximantes. Sin embargo, resultó a fines del siglo antepasado que tal procedimiento no es razonable; por ejemplo H.A. Schwarz demostró que para una superficie tan elemental como el área curvada de un cilindro, el “contenido” depende de la aproximación por triángulos de tal manera que cualquier número positivo e incluso infinito puede resultar como “contenido límite” de la aproximación. Tendremos que elegir, entonces, otro planteo para definir el contenido de una superficie, el cual, sin embargo, parecerá igualmente natural.

Para motivar la definición del contenido de una superficie en $\mathbb{R}^3$, consideremos una superficie F con la parametrización (f, I^0) , donde I es el intervalo cerrado

$$I = \{(u, v) \mid a \leq u \leq b, c \leq v \leq d\},$$

e I^0 es el intervalo abierto correspondiente. Sea

$$P = \{I_{k,l}\}_{\substack{k=1,\dots,m, \\ l=1,\dots,n}}$$

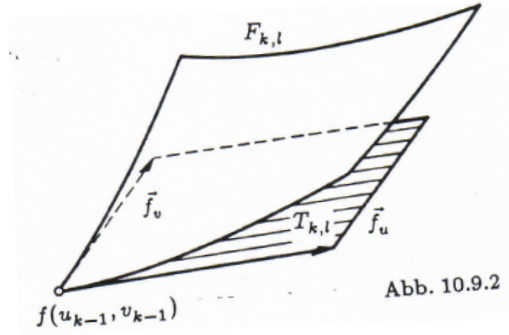
una partición de I , donde

$$I_{k,l} = \{(u, v) \mid u_{k-1} \leq u \leq u_k, v_{l-1} \leq v \leq v_l\}.$$

Esta partición define una descomposición de F en trozos de superficie, ver Figura 6.15.

Consideremos ahora el intervalo $I_{k,l}$ y el trozo de superficie correspondiente $F_{k,l} \subset F$. Ahora, si reemplazamos $F_{k,l}$ por el paralelograma $T_{k,l}$ del plano tangencial en el punto $f(u_{k-1}, v_{k-1})$, obtenemos una aproximación al contenido $|F_{k,l}|$ de $F_{k,l}$ que deseamos definir. Según la definición del plano tangencial, este paralelograma está dado por los vectores $\vec{f}_u(u_{k-1}, v_{l-1})(u_k - u_{k-1})$ y $\vec{f}_v(u_{k-1}, v_{l-1})(v_l - v_{l-1})$; su contenido es

$$\|\vec{f}_u(u_{k-1}, v_{l-1}) \times \vec{f}_v(u_{k-1}, v_{l-1})\| (u_k - u_{k-1})(v_l - v_{l-1}),$$

FIGURA 6.16. El trozo de superficie $F_{k,l}$.

por lo tanto una aproximación apropiada del contenido de superficie de F sería dada por la suma de Riemann

$$\sum_{k,l} \|\vec{f}_u(u_{k-1}, v_{l-1}) \times \vec{f}_v(u_{k-1}, v_{l-1})\| (u_k - u_{k-1})(v_l - v_{l-1}),$$

la cual converge bajo hipótesis apropiadas a la integral

$$\iint_I \|\vec{f}_u(u, v) \times \vec{f}_v(u, v)\| \, d(u, v).$$

Estas consideraciones motivan la siguiente definición, en la cual reemplazamos I por dominios más generales S .

Definición 6.18. Sea F una superficie en $\mathbb{R}^3$ con la parametrización (f, S) . Si existe la integral

$$\iint_S \|\vec{f}_u(u, v) \times \vec{f}_v(u, v)\| \, d(u, v) = \iint_F d\sigma,$$

entonces su valor se llama contenido de superficie de F .

Puesto que para una superficie F los vectores

$$\vec{f}_u(u, v) = \left\{ \frac{\partial x}{\partial u}(u, v), \frac{\partial y}{\partial u}(u, v), \frac{\partial z}{\partial u}(u, v) \right\}, \quad \vec{f}_v(u, v) = \left\{ \frac{\partial x}{\partial v}(u, v), \frac{\partial y}{\partial v}(u, v), \frac{\partial z}{\partial v}(u, v) \right\}$$

son funciones continuas y por lo tanto también $\|\vec{f}_u(u, v) \times \vec{f}_v(u, v)\|$ es continua, las integrales mencionadas arriba seguramente existen si el conjunto $S \subset \mathbb{R}^2$ es medible.

Ejemplo 6.21. Consideremos el toro dado para $R > 1$ por

$$\begin{aligned} x &= (R - \cos v) \cos u, & -\pi &\leq u \leq \pi, \\ y &= (R - \cos v) \sin u, & -\pi &\leq v \leq \pi, \\ z &= \sin v, \end{aligned}$$

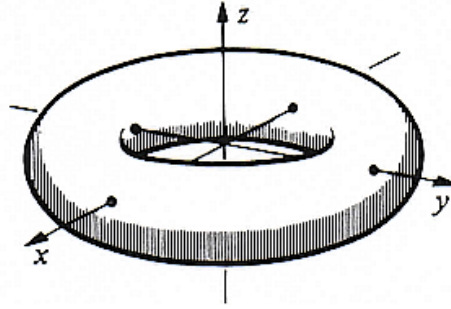


FIGURA 6.17. El toro del Ejemplo 6.21.

ver Figura 6.17, y sea F aquella parte de la superficie total del toro que corresponde a la parametrización (f, S) con

$$\begin{aligned} f(u, v) &= ((R - \cos v) \cos u, (R - \cos v) \sin u, \sin v), \\ S &= \{(u, v) \mid -\pi/2 < u < \pi/2, -\pi/2 < v < \pi/2\}. \end{aligned}$$

Queremos ahora calcular la superficie de F . Para tal efecto calculamos las cantidades

$$\begin{aligned} \vec{f}_u &= \{-(R - \cos v) \sin u, (R - \cos v) \cos u, 0\}, \\ \vec{f}_v &= \{\sin v \cos u, \sin v \sin u, \cos v\}, \\ \vec{f}_u \times \vec{f}_v &= (R - \cos v) \{\cos u \cos v, \sin u \cos v, -\sin v\}, \\ \|\vec{f}_u \times \vec{f}_v\| &= (R - \cos v) \sqrt{\cos^2 u \cos^2 v + \sin^2 u \cos^2 v + \sin^2 v} \\ &= R - \cos v, \end{aligned}$$

por lo tanto obtenemos el siguiente contenido O de nuestra superficie:

$$\begin{aligned} O &= \iint_F d\sigma = \iint_S (R - \cos v) d(u, v) = \int_{-\pi/2}^{\pi/2} \left(\int_{-\pi/2}^{\pi/2} (R - \cos v) dv \right) du \\ &= \int_{-\pi/2}^{\pi/2} (R\pi - 2) du = R\pi^2 - 2\pi. \end{aligned}$$

Comentamos que si la superficie está dada en forma explícita $z = z(x, y)$ para $(x, y) \in S$, es decir si una parametrización de F está dada por $x = u$, $y = v$ y $z = z(u, v)$, obtenemos los vectores

$$\vec{f}_u = \left\{ 1, 0, \frac{\partial z}{\partial u} \right\}, \quad \vec{f}_v = \left\{ 0, 1, \frac{\partial z}{\partial v} \right\}, \quad \vec{f}_u \times \vec{f}_v = \left\{ -\frac{\partial z}{\partial u}, -\frac{\partial z}{\partial v}, 1 \right\}$$

y por lo tanto en este caso el contenido de superficie O de F está dado por

$$O = \iint_F d\sigma = \iint_S \|\vec{f}_u \times \vec{f}_v\| d(u, v) = \iint_S \sqrt{1 + \left(\frac{\partial z}{\partial u}\right)^2 + \left(\frac{\partial z}{\partial v}\right)^2} d(u, v)$$

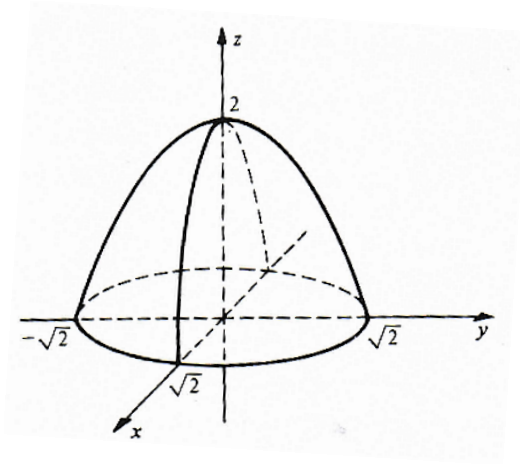


FIGURA 6.18. El paraboloides del Ejemplo 6.22.

$$= \iint_S \sqrt{1 + \left(\frac{\partial z}{\partial x}\right)^2 + \left(\frac{\partial z}{\partial y}\right)^2} d(x, y).$$

Ejemplo 6.22. Queremos calcular el contenido de superficie O de aquella parte de un paraboloides F que explícitamente está dado por

$$z = z(x, y) = 2 - x^2 - y^2, \quad 0 \leq x^2 + y^2 < 2,$$

ver Figura 6.18. Para tal efecto calculamos aquí las derivadas parciales

$$\frac{\partial z}{\partial x} = -2x, \quad \frac{\partial z}{\partial y} = -2y,$$

por lo tanto

$$O = \iint_F d\sigma = \iint_S \sqrt{1 + 4x^2 + 4y^2} d(x, y).$$

Utilizando coordenadas planas $x = r \cos \varphi$, $y = r \sin \varphi$ obtenemos

$$\begin{aligned} O &= \int_0^{2\pi} \left(\int_0^{\sqrt{2}} \sqrt{1 + 4r^2 \cos^2 \varphi + 4r^2 \sin^2 \varphi} r dr \right) d\varphi \\ &= \int_0^{2\pi} \left(\int_0^{\sqrt{2}} \sqrt{1 + 4r^2} r dr \right) d\varphi = \int_0^{2\pi} \left[\frac{1}{12} (1 + 4r^2)^{3/2} \right]_{r=0}^{r=\sqrt{2}} d\varphi = \frac{13}{6} \int_0^{2\pi} d\varphi = \frac{13}{3} \pi. \end{aligned}$$

Tal como definimos el contenido de una superficie podemos proceder para definir las *integrales de superficie*. Tales integrales aparecen en aplicaciones de física, en la mecánica del medio continuo, etc.

Sea F una superficie en $\mathbb{R}^3$ dada y sea $G : \mathbb{R}^3 \rightarrow \mathbb{R}$ una función definida sobre F . Preguntamos ¿cómo podemos integrar la función G sobre la superficie F ? Un procedimiento

bastante obvio consiste en seleccionar en cada trozo de superficie $F_{k,l}$ un punto, por ejemplo $f(u_{k-1}, v_{l-1})$, formar la suma

$$\sum_{k,l} G(f(u_{k-1}, v_{l-1})) |F_{k,l}|,$$

y considerar su límite bajo refinamiento. Reemplazando en esta sumatoria $|F_{k,l}|$ por la aproximación $|T_{k,l}|$ obtenemos la suma de Riemann

$$\sum_{k,l} G(f(u_{k-1}, v_{l-1})) \|\vec{f}_u(u_{k-1}, v_{l-1}) \times \vec{f}_v(u_{k-1}, v_{l-1})\| (u_k - u_{k-1})(v_l - v_{l-1}),$$

la cual converge bajo hipótesis apropiadas a la integral

$$\iint_I G(f(u, v)) \|\vec{f}_u(u, v) \times \vec{f}_v(u, v)\| d(u, v).$$

Estas consideraciones motivan la siguiente definición de la integral de superficie.

Definición 6.19. Sea F una superficie en $\mathbb{R}^3$ con la parametrización (f, S) ; además sea $G : \mathbb{R}^3 \rightarrow \mathbb{R}$ una función continua sobre F . Si la integral

$$\iint_S G(f(u, v)) \|\vec{f}_u(u, v) \times \vec{f}_v(u, v)\| d(u, v) = \iint_F G d\sigma$$

existe, entonces su valor se llama integral de superficie de G sobre F .

Ejemplo 6.23. Queremos calcular la integral

$$\iint_F G d\sigma,$$

donde F es la superficie del paraboloide del Ejemplo 6.22 y $G(x, y) := x^2 + y^2$. Tal como en el Ejemplo 6.22 obtenemos aquí

$$\begin{aligned} \iint_F G d\sigma &= \iint_S (u^2 + v^2) \sqrt{1 + 4u^2 + 4v^2} d(u, v) \\ &= \int_0^{\sqrt{2}} \left(\int_0^{2\pi} (r^2 \cos^2 \varphi + r^2 \sin^2 \varphi) r \sqrt{1 + 4r^2 \cos^2 \varphi + 4r^2 \sin^2 \varphi} d\varphi \right) dr \\ &= 2\pi \int_0^{\sqrt{2}} r^3 \sqrt{1 + 4r^2} dr = \frac{\pi}{8} \int_1^3 (\varrho^2 - 1) \varrho^2 d\varrho = \frac{149}{30} \pi. \end{aligned}$$

6.10. El Teorema Integral de Green

Teorema 6.18 (Teorema Integral de Green). Sea $\vec{V} = (L, M) : \mathbb{R}^2 \rightarrow \mathbb{V}^2$ un campo vectorial. Sea $\vec{V}$ definido sobre un conjunto abierto $D \subset \mathbb{R}^2$, y sea $\vec{V} \in C^1(D)$. Sea S un recinto

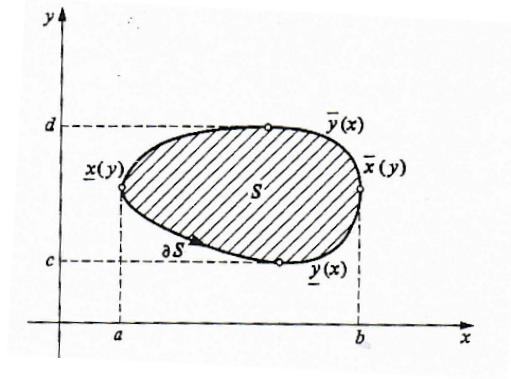


FIGURA 6.19. Ilustración de la demostración del Teorema 6.18.

estándar con $\bar{S} \subset D$ y una frontera ∂D suave por trozos. Sea ∂S orientada de manera positiva, es decir, sea el interior de S siempre ubicado a la izquierda. Entonces se tiene que

$$\iint_S \left(\frac{\partial L}{\partial y} - \frac{\partial M}{\partial x} \right) d(x, y) = - \int_{\partial S} (L dx + M dy).$$

Demostración. Puesto que S es un recinto estándar, existen intervalos $[a, b]$ y $[c, d]$ y funciones $\underline{y}(x)$ e $\bar{y}(x)$ continuas sobre $[a, b]$ y funciones $\underline{x}(y)$ y $\bar{x}(y)$ continuas sobre $[c, d]$ tales que

$$\bar{S} = \bigcup_{x \in [a, b]} \{(x, y) \mid \underline{y}(x) \leq y \leq \bar{y}(x)\} = \bigcup_{y \in [c, d]} \{(x, y) \mid \underline{x}(y) \leq x \leq \bar{x}(y)\}.$$

1. Aprovecharemos ahora que S es proyectable en la dirección del eje y .

a) Parametrizamos ∂S de la siguiente manera:

$$x(t) = \begin{cases} a + t(b - a) & \text{para } t \in [0, 1], \\ b & \text{para } t \in [1, 2], \\ b - (t - 2)(b - a) & \text{para } t \in [2, 3], \\ a & \text{para } t \in [3, 4], \end{cases}$$

$$y(t) = \begin{cases} \underline{y}(a + t(b - a)) & \text{para } t \in [0, 1], \\ \underline{y}(b) + (t - 1)(\bar{y}(b) - \underline{y}(b)) & \text{para } t \in [1, 2], \\ \bar{y}(b - (t - 2)(b - a)) & \text{para } t \in [2, 3], \\ \bar{y}(a) - (t - 3)(\bar{y}(a) - \underline{y}(a)) & \text{para } t \in [3, 4]. \end{cases}$$

b) Podemos ahora calcular la integral

$$\iint_S \frac{\partial L}{\partial y} d(x, y)$$

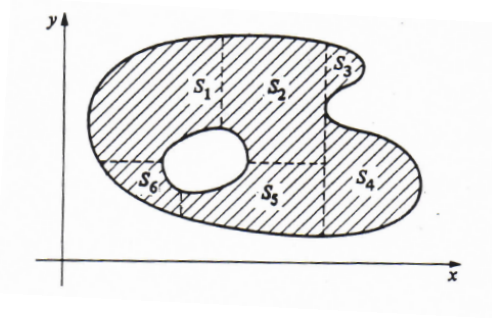


FIGURA 6.20. Una región de Green (ver Definición 6.20).

como integral iterada en la siguiente forma:

$$\begin{aligned}
 \iint_S \frac{\partial L}{\partial y} d(x, y) &= \int_a^b \left(\int_{\underline{y}(x)}^{\bar{y}(x)} \frac{\partial L}{\partial y} dy \right) dx \\
 &= \int_a^b \left(L(x, \bar{y}(x)) - L(x, \underline{y}(x)) \right) dx \\
 &= - \int_b^a L(x, \bar{y}(x)) dx - \int_a^b L(x, \underline{y}(x)) dx \\
 &= - \int_0^4 L(x(t), y(t)) x'(t) dt = - \int_{\partial S} L dx.
 \end{aligned}$$

Aquí utilizamos la substitución $x = x(t) = a + t(b - 1)$ para $t \in [0, 1]$ para deducir que

$$\int_a^b L(x, \underline{y}(x)) dx = \int_0^1 L(x(t), y(t)) x'(t) dt,$$

mientras que la substitución $x = x(t) = b - (t - 2)(b - a)$ para $t \in [2, 3]$ nos entrega que

$$- \int_b^a L(x, \bar{y}(x)) dx = - \int_2^3 L(x(t), y(t)) x'(t) dt,$$

y los integrales sobre $[1, 2]$ y $[3, 4]$ poseen el valor cero.

2. Analogamente se establece que

$$\iint_S \left(-\frac{\partial M}{\partial x} \right) d(x, y) = - \int_{\partial S} M dy.$$

Esto concluye la demostración del teorema. ■

El teorema sigue válido bajo presuposiciones más débiles. Por ejemplo, basta suponer que el conjunto S es una *región de Green*, ver Figura 6.20.

Definición 6.20. Un conjunto de puntos $S \subset \mathbb{R}^n$ se llama región de Green si S puede ser descompuesto en un número finito de recintos estándar $S_1, \dots, S_N$ con fronteras suaves por trozos.

Para ver que el Teorema Integral de Green sigue siendo válido para regiones de Green, siempre que consideremos la orientación positiva de ∂S , aplicamos el teorema a cada recinto estándar S_i , $i = 1, \dots, N$:

$$\iint_{S_i} \left(\frac{\partial L}{\partial y} - \frac{\partial M}{\partial x} \right) d(x, y) = - \int_{\partial S_i} (L dx + M dy).$$

Sumando sobre $i = 1, \dots, N$ obtenemos

$$\iint_S \left(\frac{\partial L}{\partial y} - \frac{\partial M}{\partial x} \right) d(x, y) = - \int_{\partial S} (L dx + M dy),$$

puesto que cada trozo de frontera interior del lado derecho aparece dos veces en sentidos opuestos, lo que significa que las contribuciones correspondientes se cancelan.

Ejemplo 6.24. Sea $S = \{(x, y) \mid 0 \leq x^2 + y^2 \leq 1\}$, entonces según el Teorema de Green se tiene que

$$\int_{\partial S} (2xy^3 dx + 3x^2y^2 dy) = - \iint_S (6xy^2 - 6xy^2) d(x, y) = 0.$$

Ejemplo 6.25. Sea $S = \{(x, y) \mid 0 \leq x^2 + y^2 \leq 1\}$. Queremos calcular

$$J = \int_{\partial S} ((x^4 - y^3) dx + (x^3 - y^4) dy).$$

Según el Teorema de Green se tiene que

$$J = - \iint_S (-3y^2 - 3x^2) d(x, y) = 3 \iint_S (x^2 + y^2) d(x, y).$$

Introduciendo coordenadas polares $x = r \cos \varphi$, $y = r \sin \varphi$ obtenemos

$$J = 3 \int_0^1 \left(\int_0^{2\pi} r^2 \cdot r d\varphi \right) dr = 6\pi \int_0^1 r^3 dr = \frac{3\pi}{2}.$$

6.11. El Teorema Integral de Gauss

Teorema 6.19 (Teorema Integral de Gauss). Sea $\vec{V} : \mathbb{R}^3 \rightarrow \mathbb{V}^3$ un campo vectorial definido sobre el conjunto abierto $D \subset \mathbb{R}^3$; sea $\vec{V} \in C^1(D)$. Sea S un recinto estándar con $\bar{S} \subset D$, y sea ∂S un subconjunto de un número finito de superficies. Si denotamos por $\vec{n}$ el vector normal exterior sobre ∂S , entonces se tiene que

$$\iiint_S \operatorname{div} \vec{V} d(x, y, z) = \iint_{\partial S} (\vec{V} \cdot \vec{n}) d\sigma.$$

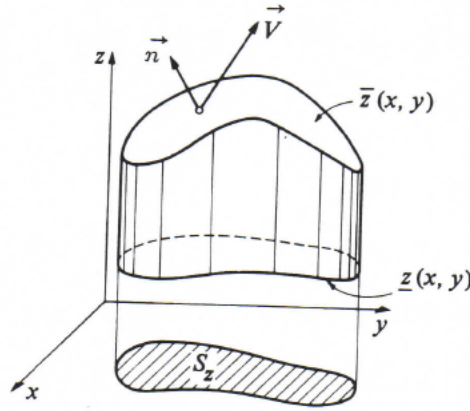


FIGURA 6.21. Ilustración de la demostración del Teorema 6.19.

Demostración. Sean $\vec{V} = (L, M, N)$ y $\vec{n} = (\cos \alpha, \cos \beta, \cos \gamma)$.

1. Puesto que S es proyectable con respecto al eje z , obtenemos

$$\begin{aligned}
 & \iiint_S \frac{\partial N}{\partial z}(x, y, z) \, d(x, y, z) \\
 &= \iint_{S_z} \left(\int_{\underline{z}(x, y)}^{\bar{z}(x, y)} \frac{\partial N}{\partial z}(x, y, z) \, dz \right) d(x, y) \\
 &= \iint_{S_z} N(x, y, \bar{z}(x, y)) \, d(x, y) - \iint_{S_z} N(x, y, \underline{z}(x, y)) \, d(x, y) \\
 &= \iint_{S_z} N(x, y, \bar{z}(x, y)) \frac{\sqrt{\left(\frac{\partial \bar{z}}{\partial x}(x, y)\right)^2 + \left(\frac{\partial \bar{z}}{\partial y}(x, y)\right)^2 + 1}}{\sqrt{\left(\frac{\partial \bar{z}}{\partial x}(x, y)\right)^2 + \left(\frac{\partial \bar{z}}{\partial y}(x, y)\right)^2 + 1}} d(x, y) \\
 &\quad + \iint_{S_z} N(x, y, \underline{z}(x, y)) \frac{-\sqrt{\left(\frac{\partial \underline{z}}{\partial x}(x, y)\right)^2 + \left(\frac{\partial \underline{z}}{\partial y}(x, y)\right)^2 + 1}}{\sqrt{\left(\frac{\partial \underline{z}}{\partial x}(x, y)\right)^2 + \left(\frac{\partial \underline{z}}{\partial y}(x, y)\right)^2 + 1}} d(x, y),
 \end{aligned}$$

ver Figura 6.21. Ahora, considerando que sobre la superficie límite superior (el “techo” de S) se tiene que

$$\frac{1}{\sqrt{\left(\frac{\partial \bar{z}}{\partial x}(x, y)\right)^2 + \left(\frac{\partial \bar{z}}{\partial y}(x, y)\right)^2 + 1}} = \cos \gamma$$

mientras que sobre la superficie límite inferior (el “piso” de S),

$$\frac{-1}{\sqrt{\left(\frac{\partial z}{\partial x}(x, y)\right)^2 + \left(\frac{\partial z}{\partial y}(x, y)\right)^2 + 1}} = \cos \gamma$$

y que $\cos \gamma = 0$ para posibles trozos de superficie paralelos al eje z , llegamos a

$$\iiint_S \frac{\partial N}{\partial z}(x, y, z) \, d(x, y, z) = \iint_{\partial S} N \cos \gamma \, d\sigma.$$

2. Aprovechando que S también es proyectable con respecto a los ejes y y x , podemos establecer análogamente que

$$\begin{aligned} \iiint_S \frac{\partial M}{\partial y}(x, y, z) \, d(x, y, z) &= \iint_{\partial S} M \cos \beta \, d\sigma, \\ \iiint_S \frac{\partial L}{\partial x}(x, y, z) \, d(x, y, z) &= \iint_{\partial S} L \cos \alpha \, d\sigma. \end{aligned}$$

3. Sumando y considerando que

$$\operatorname{div} \vec{V} = \frac{\partial L}{\partial x} + \frac{\partial M}{\partial y} + \frac{\partial N}{\partial z} \quad \text{y} \quad L \cos \alpha + M \cos \beta + N \cos \gamma = \vec{V} \cdot \vec{n}$$

podemos concluir la demostración del teorema. ■

Ejemplo 6.26. Sean

$$\begin{aligned} S &= \left\{ (x, y, z) \mid 0 \leq x^2 + y^2 \leq 1, 0 \leq z \leq \sqrt{1 - (x^2 + y^2)} \right\}, \\ \vec{V} &= \{xz^2, x^2y - z^3, 2xy + y^2z\}. \end{aligned}$$

Queremos calcular la integral

$$J = \iint_{\partial S} (\vec{V} \cdot \vec{n}) \, d\sigma.$$

Según el Teorema de Gauss obtenemos

$$\begin{aligned} \iint_{\partial S} (\vec{V} \cdot \vec{n}) \, d\sigma &= \iiint_S \operatorname{div} \vec{V} \, d(x, y, z) \\ &= \iiint_S \left(\frac{\partial}{\partial x}(xz^2) + \frac{\partial}{\partial y}(x^2y - z^3) + \frac{\partial}{\partial z}(2xy + y^2z) \right) d(x, y, z) \\ &= \iiint_S (x^2 + y^2 + z^2) \, d(x, y, z). \end{aligned}$$

Resolviendo la última integral utilizando coordenadas polares esféricas obtenemos

$$J = \frac{2}{5}\pi.$$

6.12. El Teorema Integral de Stokes

Teorema 6.20 (Teorema Integral de Stokes). Sea $\vec{V} = (L, M, N) : \mathbb{R}^3 \rightarrow \mathbb{V}^3$ un campo vectorial; sea $\vec{V}$ definido sobre un conjunto abierto $D \subset \mathbb{R}^3$, y sea $\vec{V} \in C^1(D)$. Sea F una superficie con la parametrización (f, S) contenida en D , y sea $f \in C^2(S)$. Además, sea $G \subset F$ un conjunto con la imagen recíproca $\tilde{G} \subset S$, donde $\tilde{G}$ sea una región de Green. En este caso se tiene que

$$\iint_G (\text{rot } \vec{V} \cdot \vec{n}) \, d\sigma = \int_{\partial G} (L \, dx + M \, dy + N \, dz).$$

Demostración.

1. Sea $(\tilde{k} = (\tilde{u}(t), \tilde{v}(t)), [a, b])$ una parametrización de $\partial\tilde{G}$. Si definimos

$$\tilde{L}(u, v) := L(x(u, v), y(u, v), z(u, v)),$$

entonces según el Teorema de Green se tiene que

$$\begin{aligned} \int_{\partial G} L \, d\sigma &= \int_a^b \tilde{L}(\tilde{u}(t), \tilde{v}(t)) \frac{d}{dt} (x(\tilde{u}(t), \tilde{v}(t))) \, dt \\ &= \int_a^b \tilde{L}(\tilde{u}(t), \tilde{v}(t)) \left(\frac{\partial x}{\partial u}(\tilde{u}(t), \tilde{v}(t)) \tilde{u}'(t) + \frac{\partial x}{\partial v}(\tilde{u}(t), \tilde{v}(t)) \tilde{v}'(t) \right) dt \\ &= \iint_{\tilde{G}} \left[\frac{\partial}{\partial u} \left(\tilde{L}(u, v) \frac{\partial x}{\partial v}(u, v) \right) - \frac{\partial}{\partial v} \left(\tilde{L}(u, v) \frac{\partial x}{\partial u}(u, v) \right) \right] d(u, v). \end{aligned}$$

Según el Teorema de Scharz sabemos que

$$\frac{\partial^2 x}{\partial u \partial v} = \frac{\partial^2 x}{\partial v \partial u}$$

y por lo tanto

$$\begin{aligned} &\left[\frac{\partial}{\partial u} \left(\tilde{L}(u, v) \frac{\partial x}{\partial v}(u, v) \right) - \frac{\partial}{\partial v} \left(\tilde{L}(u, v) \frac{\partial x}{\partial u}(u, v) \right) \right] \\ &= \left(\frac{\partial L}{\partial x} \frac{\partial x}{\partial u} + \frac{\partial L}{\partial y} \frac{\partial y}{\partial u} + \frac{\partial L}{\partial z} \frac{\partial z}{\partial u} \right) \frac{\partial x}{\partial v} + \tilde{L} \cdot \frac{\partial^2 x}{\partial v \partial u} \\ &\quad - \left(\frac{\partial L}{\partial x} \frac{\partial x}{\partial v} + \frac{\partial L}{\partial y} \frac{\partial y}{\partial v} + \frac{\partial L}{\partial z} \frac{\partial z}{\partial v} \right) \frac{\partial x}{\partial u} - \tilde{L} \frac{\partial^2 x}{\partial u \partial v} \\ &= - \left(\frac{\partial x}{\partial u} \frac{\partial y}{\partial v} - \frac{\partial x}{\partial v} \frac{\partial y}{\partial u} \right) \frac{\partial L}{\partial y} + \left(\frac{\partial x}{\partial v} \frac{\partial z}{\partial u} - \frac{\partial x}{\partial u} \frac{\partial z}{\partial v} \right) \frac{\partial L}{\partial z} \\ &= -D_z \frac{\partial L}{\partial y} + D_y \frac{\partial L}{\partial z}, \end{aligned}$$

utilizando la notación de la Sección 6.8. Entonces obtenemos

$$\begin{aligned}\int_{\partial G} L \, dx &= \iint_{\tilde{G}} \left(\frac{\partial L}{\partial y} \frac{D_z}{D} + \frac{\partial L}{\partial z} \frac{D_y}{D} \right) D \, d(u, v) \\ &= \iint_G \left(-\frac{\partial L}{\partial y} \cos \gamma + \frac{\partial L}{\partial z} \cos \beta \right) d\sigma.\end{aligned}$$

2. Análogamente obtenemos las ecuaciones

$$\begin{aligned}\int_{\partial G} M \, dy &= \iint_G \left(-\frac{\partial M}{\partial z} \cos \alpha + \frac{\partial M}{\partial x} \cos \gamma \right) d\sigma, \\ \int_{\partial G} N \, dz &= \iint_G \left(-\frac{\partial N}{\partial x} \cos \beta + \frac{\partial N}{\partial y} \cos \alpha \right) d\sigma.\end{aligned}$$

3. Sumando y considerando que

$$\left(\frac{\partial N}{\partial y} - \frac{\partial M}{\partial z} \right) \cos \alpha + \left(\frac{\partial L}{\partial z} - \frac{\partial N}{\partial x} \right) \cos \beta + \left(\frac{\partial M}{\partial x} - \frac{\partial L}{\partial y} \right) \cos \gamma = \operatorname{rot} \vec{V} \cdot \vec{n}$$

llegamos al enunciado del teorema. ■

Ejemplo 6.27. Consideremos el campo vectorial $\vec{V} = \{-y, x, 1\}$ y como superficie F el hemisferio superior dado por $f = (x, y, z = \sqrt{4 - x^2 - y^2})$, $S = \{(x, y) \mid x^2 + y^2 < 4\}$. Sea $G = F$; entonces ∂G es el círculo en el plano (x, y) con el origen como centro y el radio 2. Inmediatamente calculamos aquí $\operatorname{rot} \vec{V} = \{0, 0, 2\}$ y el vector normal

$$\vec{n} = \frac{\left\{ -\frac{\partial z}{\partial x}, -\frac{\partial z}{\partial y}, 1 \right\}}{\sqrt{1 + \left(\frac{\partial z}{\partial x} \right)^2 + \left(\frac{\partial z}{\partial y} \right)^2}},$$

luego obtenemos el producto escalar

$$\operatorname{rot} \vec{V} \cdot \vec{n} = \frac{2}{\sqrt{1 + \left(\frac{\partial z}{\partial x} \right)^2 + \left(\frac{\partial z}{\partial y} \right)^2}},$$

entonces

$$\iint_G \operatorname{rot} \vec{V} \cdot \vec{n} \, d\sigma = 2 \iint_{x^2 + y^2 \leq 4} d(x, y) = 8\pi$$

y por lo tanto

$$\int_{\partial G} (-y \, dx + x \, dy + dz) = 8\pi.$$